University of Sussex

Predictions for the Infrared numbercounts and star formation histories from a semi analytic model of galaxy formation.

Sorour Shamshiri

Submitted for the degree of Doctor of Philosophy

University of Sussex

September 21, 2016

Declaration

I hereby declare that this thesis has not been and will not be submitted in whole or in part to another University for the award of any other degree.

Signature:

Sorour Shamshiri

UNIVERSITY OF SUSSEX

Sorour Shamshiri, Doctor of Philosophy

Prediction of Semi analytic model. Star Formation History and Infrared Luminosity Function

Summary

One of the most fundamental probes of the physics that underpins galaxy evolution is the star formation rate (SFR) as a function of cosmic time. In addition, the statistical properties of galaxy populations are another important key to understand how the universe has been evolving. It is known that the far-infrared emission from galaxies is strongly correlated with obscured star formation and forms a significant part of cosmic infrared background. We thus investigate the variation of the SFR of galaxies over time by comparing predictions of the L-Galaxies semi-analytic model with observations of the far infrared (FIR) luminosity and number counts.

In the first part of this thesis, we follow the star formation histories (SFHs) of galaxies and use these to construct stellar spectra in post-processing. We then contrast model SFHs from the Millennium Simulation with observed ones from the VESPA algorithm as applied to the SDSS-7 catalogue when this has been characterized by mass and colour of galaxies.

In order to investigate the SAM model prediction, I extend L-galaxies to predict far infrared fluxes and construct mock catalogues which are fed into SMAP in order to provide simulated maps. LFs have also been estimated for model galaxies at different redshifts. The results are compared with observations from Herschel. To conclude, our model underestimates the number density of galaxies at bright sources (e.g fluxes above 0.02 Jy) also

does not produce high luminosity objects especially at higher redshifts (e.g $z > 1$) . We show that by fitting the simulated IR luminosity function to observed LIR, our model is able to produce more bright sources at high redshifts and match reasonably well to the observed number counts.

Acknowledgements

Firstly, I would like to express my sincere appreciation to my supervisor Prof Seb Oliver for all the academic and friendly supports during the last years. Special thank you to my second supervisor Prof Peter Thomas for the continuous support of my Ph.D study and related research, for his patience, encouragement and immense knowledge. His guidance helped me in all the time during my time at Sussex.

Thank you to my mum and dad for their unconditional support that help me to forwards in the hardest of all times. Words can not express how grateful I am to my parents, sister, and brother for all of the sacrifices that they have made on my behalf. This journey would be impossible without having my husband beside me. He is not only my sole mate but also my best friend. My boys, Aryan and Raika, were my reasons to move on and never give up and alway look at the bright side. Thank you to my friends in Brighton. Thank you so much, Antonio, Charlotte and Samane, without you my life would not have been the same

Thanks, Mamnoon.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 State of the Art

1.1.1 ΛCDM Cosmology

The occurrence of the advance technology has revolutionised our understanding of previously unreachable Universe and enabled us to practically test the simulation breakthrough against robust observational data. As described in Baugh (2006), a combination of two factors made by spectacular breakthroughs of the powerful technology shifted the emphasis of astronomical goals towards galaxy formation. These factors are the firmer establishment of Lambda Cold Dark Matter (ΛCDM) and more statistical information of galaxy population from high redshift to the local universe (Springel et al. (2005); Frieman, Turner & Huterer (2008)). The ΛCDM model suggests that the dark matter is elementary dissipationless and collisionless particles that can interact gravitationally and it cannot be cooled by radiative processes. The model also assumes a flat, homogeneous and isotropic Universe in concordance with the cosmological principle.

It predicts that structures grow hierarchically from weak primordial density perturbations seeded by inflation in the rapidly expanding early Universe. The baryons fall into these gravitational instabilities and eventually turn into the rich structure that we see all over the present Universe. Acoustic oscillations raised by these primordial fluctuations leave an imprint in the matter power spectrum. The triumphs of the model have been vastly enhanced by recent observations. The precise measurement of the temperature anisotropies in cosmic microwave background (CMB) is the first robust confirmation of the model (Efstathiou (2002); Jones et al. (2006)). The initial confirmation of these anisotropies has been made by the COBE satellite (Smoot et al., 1992) and their power spectrum has been displayed since (Hanany et al. (2000); Percival et al. (2001); Tegmark et al. (2004);

Hinshaw et al. (2007)). The compelling support of the model has continued by detection of three Doppler peaks caused by the baryon acoustic oscillations in the photon-baryon fluid just before recombination (Hanany et al. (2000); Jones et al. (2006); Hinshaw et al. (2007). The advanced technology not only provides the observational opportunity to test the lambda CDM model, but is also utilised to study galaxy formation and evolution.

1.1.2 Cosmological Simulations

The foundation of a cosmological paradigm with tremendous predictive power and the present improvement in observations are working together towards a better understanding of galaxy formation and evolution. However, we require a realistic leading theory of galaxy formation in order to have a better understanding of the physical processes that construct galaxy properties. Powerful cosmological simulations take initial conditions and compute the properties of dark matter and baryons as they form structures hierarchically. Over the last decade, advanced robust computer hardware and computational capabilities have enabled direct N-body simulations to provide a precise image of formation and evolution of dark matter from primordial gravitational perturbations to the formation of galaxies as a function of redshift (The Millennium Run, Springel et al. (2005)). The predicted cosmic web is presented in Fig. 1.1. The image illustrates the projected density field in a slice of thickness $15h^{-1}$Mpc, zoomed in by a factor of four with respect to the previous panels. Although the physics behind galaxy formation and evolution is still poorly understood in detail, high-resolution N$-$body simulations have provided a unique impact on our knowledge on the subject.

1.1.3 Semi analytic model

Tracking the evolution of baryons in a combined simulation of gas and dark matter is a complex operation since it involves physical processes such as feedback and star formation, that are not completely understood yet. Therefore, galaxy formation is a complicated procedure and we need a combination of numerical simulations and analytical models and observations in order to understand the whole process. The Semi analytic models of galaxy formation, first introduced by White & Rees (1978), are known as the best analytical method of generating an accurate prediction of galaxy properties and providing a straightforward contrast with observed surveys. It tracks the formation of dark matter halos caused by gravitational instabilities while the evolution of baryonic components is calculated by using a series of simple physical procedures. Dark matter struc-

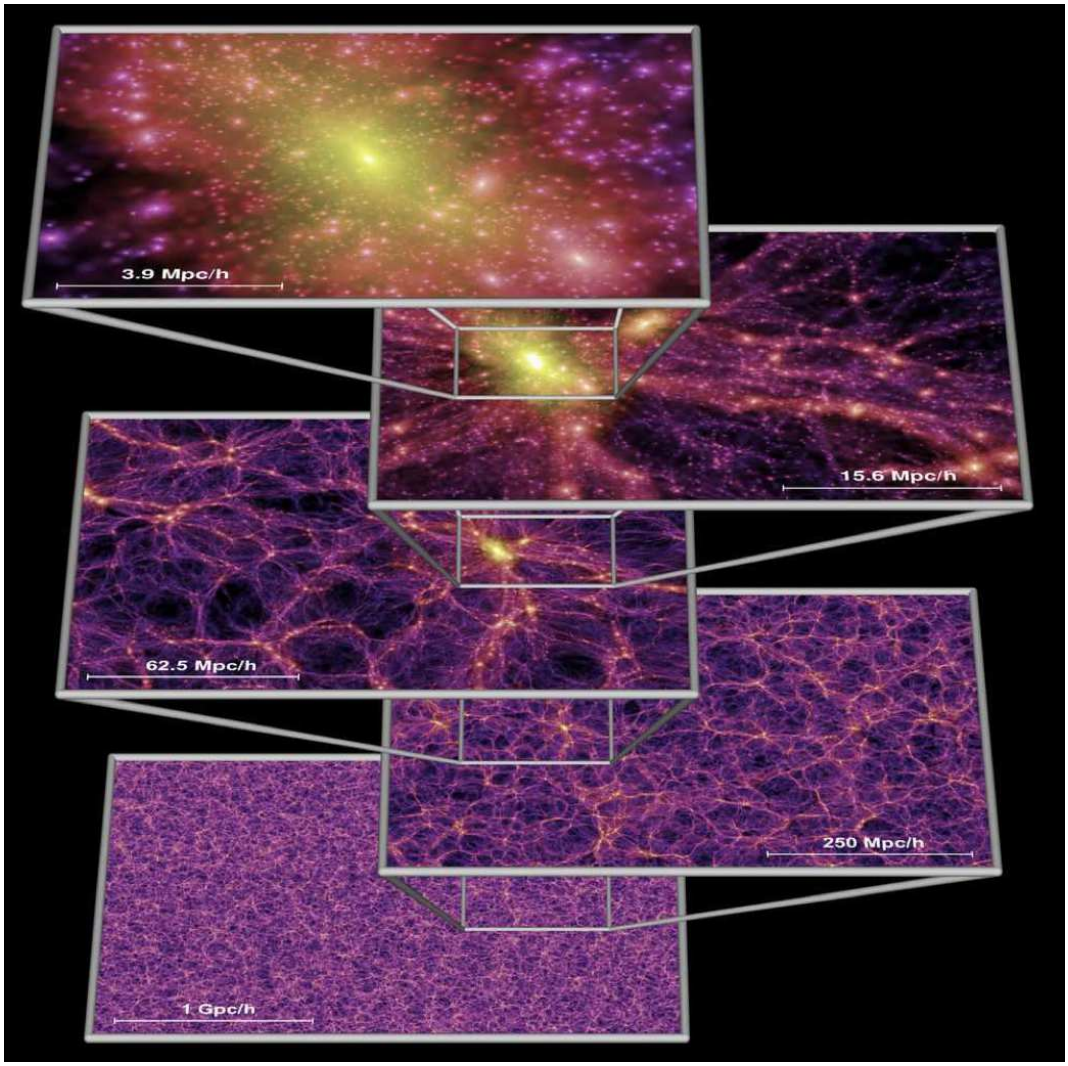

Figure 1.1: The dark matter distribution from the Millennium Simulation. The image shows the projected density field in a slice of thickness $15h^{-1}$Mpc at redshift zero. White squares indicate the regions zoomed in by factors of 4 in each case (Springel et al., 2005).

tures are studied by either using N−body techniques, Monte Carlo, or a Press-Schechter (Press & Schechter, 1974). One of the advantages is providing a comprehensive picture of galaxy evolution at a relatively small computational cost. In addition, SA models provide a wide range of predictions for observable properties of galaxies ranging from isolated dwarf galaxies to the most massive ones, from high redshifts to the local Universe. For this reason, and considering the flexibility of the model, they can be a crucial means of elaboration on observable outcomes of current cosmological theories and the interpretation of observations, mainly at high redshifts.

1.2 The Munich Semi Analytic Model

The work presented in this thesis was developed using the latest version of the L-Galaxies semi-analytic model (Henriques et al. (2015), thereafter HWT15). Over the last decades, the physical processes that shaped galaxy formation and its physical foundation have improved followed by updated knowledge gained from observational data sets. The outcome of the model is primarily governed by the underlying distribution of dark matter as it controls the mass assembly of galaxies. Therefore, in order to understand the structure formation, it is essential to use parameterised equations on top of an N body simulation encompassing a volume that not only resolves big structures, such as massive clusters, but also recovers individual dwarf galaxies.

The Munich SA model has been constructed for two simulations of the evolution of dark matter structure, namely Millennium (MR, Springel et al., 2005) and Millennium-II (MR2, Boylan-Kolchin et al., 2009). The latter is 5 times smaller than MR but 125 times greater in respects of mass resolution. Both MR and MR2 include the same number of particles and produce data at 58 snapshots separated in time by about $3{\times}10^8$ yr at low redshift (less at earlier times).

As described in Henriques et al. (2015), the L-Galaxies model adopts the hierarchical growth of structure from the Millennium Simulation (MR, Springel et al., 2005) and scaled it using the method of Angulo & White (2010) to the *Planck* cosmology (Planck Collaboration XVI 2014): $\Omega_{\rm m} = 0.315$, $\Omega_\Lambda = 0.683$, $\Omega_{\rm b} = 0.0488$, $n_{\rm s} = 0.958$, $\sigma_8 = 0.826$, $h = 0.673$. This then gives a box size of $480.3\,h^{-1}$ Mpc and a particle mass of $9.61 \times 10^8 h^{-1}$ $\rm M_\odot$.

Table 1.1: Estimations for the free parameters controlling galaxy formation physics as described in Henriques et al. (2015)

parameter	description	value	unit
$\alpha_{\rm SF}$	(SF eff)	0.030	
$\sum_{\rm SF}$	Gas density threshold	0.26	$10^{10} M_\odot\ pc^{-2}$
$\alpha_{\rm SF,burst}$	SF burst eff	0.72	
$\beta_{\rm burst}$	SF burst slope	2.0	
$k_{\rm AGN}$	Radio feedback eff	4.4×10^3	$M_\odot\ yr^{-1}$
$f_{\rm BH}$	BH growth eff	0.036	
$V_{\rm BH}$	Quasar growth scale	730	km s^{-1}
ϵ	Mass-loading eff	2.6	
$V_{\rm reheat}$	Mass-loading scale	430	km s^{-1}
β_1	Mass-loading slope	0.80	
μ	SN ejection eff	0.64	
V_{eject}	SN ejection scale	110	km s^{-1}
β_2	SN ejection slope	0.81	
λ	Ejecta reincorporation	3.6×10^{10}	yr
$M_{r.p}$	Ram-pressure threshold	1.2×10^4	$10^{10} M_\odot$
$R_{\rm merger}$	Major-merger threshold	0.41	
$\alpha_{\rm friction}$	Dynamical friction	2.6	
y	Metal yield	0.045	

1.3 Physical Processes

SAM models employ parametrised equations in order to explain galaxy formation. In this thesis, I only describe in detail those physical processes that directly affect star formation (refer to Henriques et al. (2015) for full descriptions). The parameters controlling all physical processes involved in galaxy formation and evolution from the latest version of L-Galaxies are shown in Table 1.1.

1.3.1 Cooling mode

As originally introduced by the White & Frenk (1991) approach, each collapsed dark matter structure is assumed to have a mean fixed baryon fraction of $f_b^{\rm cos} = 15.5\%$ for the Planck cosmology. This infalling material is presumed to be diffuse gas and thus being

shock heated as it joins the halo.

The infalling operation is expected to be influenced by the size and redshift of DM halos at the time of their formation. For low-mass halos at early stages, where the cooling radius can be greater than the virial radius, infalling gas concentrates within the halo dynamical time and a quasi-static atmosphere can not form. At late times and for high mass halos, however the falling gas enclosed within the virial radius has had adequate time to cool. Therefore, a quasi-static hot atmosphere is created after the infalling gas has been heated by the shocks to the virial temperature and the cooling radius continues to propagate towards the virial radius of the halo. This gas would have cooled down later and its formation into the centre of the dark matter halo is modelled by cooling flows. In both scenarios, the cold gas perched in the central disk provides raw material for star formation.

The cooling time can be described as the ratio between the thermal energy density of the gas and the cooling rate per unit volume:

$$t_{\rm cool}(r) = \frac{3\mu m_H k T_{200c}}{2\rho_{\rm hot}(r)\Lambda(T_{\rm hot}, Z_{\rm hot})}, \tag{1.1}$$

where μ_H is the mean particle mass, k is the Boltzmann constant, $\rho_{\rm hot}$(r) is the hot gas density and $Z_{\rm hot}$ is the hot gas metallicity and Lambda is the equilibrium cooling function (Sutherland & Dopita, 1993). $T_{\rm hot}$ is the temperature of the hot gas which is expected to be the virial temperature of the halo expressed by $T_{200c} = 35.9$ (V_{200c} / km s^{-1})^2 K. The hot gas density as a function of radius for a simple isothermal distribution assumption is estimated through:

$$\rho_{\rm hot} = \frac{M_{\rm hot}}{4\pi R_{200c} r^2}. \tag{1.2}$$

The cooling radius is assumed to be where the cooling time equals the halo dynamical time and is given by:

$$r_{\rm cool} = \left[\frac{t_{\rm dyn,h}\Lambda(T_{\rm hot}, Z_{\rm hot})}{6\pi\mu_H k T_{200c} R_{200c}}\right]^{\frac{1}{2}}. \tag{1.3}$$

The definition of dynamical time $t_{\rm dyn}$ is adopted to be $t_{\rm dyn}$=$R_{\rm vir}$/ $V_{\rm vir}$ as in De Lucia, Kauffmann & White (2004). When $r_{\rm cool}$<$R_{\rm vir}$, it is expected that the halo is indeed in the cooling flow regime, and that the cooling rate is:

$$\dot{M}_{\rm cool} = M_{\rm hot}\frac{r_{\rm cool}}{R_{200c}}\frac{1}{t_{\rm dyn,h}}. \tag{1.4}$$

When $r_{\rm cool}$ >$R_{\rm vir}$, on the other hand, the halo is in the rapid infall regime and infalling

material accretes onto the central object in free-fall, thus on the halo dynamical time:

$$\dot{M}_{\rm cool} = \frac{M_{\rm hot}}{t_{\rm dyn,h}}. \tag{1.5}$$

This particular recipe was first introduced by Guo et al. (2011) in order to obtain a smoother gas condensation over time.

1.3.2 Star Formation

In this model stars are expected to form from the cold gas within the disk of each galaxy. The star formation rate (amount of gas that is converted into stars per unit time), is given by:

$$\dot{M}_\star = \alpha_{SF} \frac{M_{\rm gas} - M_{\rm crit}}{t_{\rm dyn,disk}}, \tag{1.6}$$

where $M_{\rm gas}$ is the total mass of cold gas, $t_{\rm dyn,disk} = R_\star/V_{\rm max}$, is the dynamical time of the star forming disk, and $M_{\rm crit}$ is a critical mass. Note that it is assumed that a fraction $R_{\rm ret}$ of total mass of stars ($M_\star$) is associated with massive and relatively short-lived stars and is instantaneously restored to the cold gas. $R_{ret} = 0.43$ is calculated from the Chabrier (2003) initial mass function. Therefore, the stellar mass of the disk is increased by $\delta M_\star$ = (1-R_{ret})$\dot{M}_\star \delta$t. The critical gas mass above which star formation is assumed to occur, $M_{\rm crit}$, is taken to be as given by Kauffmann (1996):

$$M_{\rm crit} = M_{\rm crit,0}(\frac{V_{200c}}{200 kms^{-1}})(\frac{R_{\rm gas}}{10 kpc}). \tag{1.7}$$

Although all versions of the Munich model have adopted $M_{\rm crit,0} = 3.8 \times 10^9 M_\odot$ as presented by Kauffmann et al. (1999), HWT15 and other works (Bigiel et al., 2008; Leroy et al., 2008) propose that star formation should be connected directly to a molecular gas component rather than the total amount of cold gas. The star formation threshold $M_{\rm crit}$ is treated as a free parameter in HWT15 model.

1.3.3 Merger-triggered star formation

Star formation is also triggered through bursts caused by merging galaxies. A fraction, e_{burst}, of the combined cold gas of the merging galaxies converted into stars, is presented in Somerville, Primack & Faber (2001) using the 'collisional starburst' model.

$$M_{\star,\rm burst} = \alpha_{\rm SF,burst} \left(\frac{M_1}{M_2}\right)^{\beta_{\rm SF,burst}} M_{\rm cold}, \tag{1.8}$$

where M_1 and M_2 are the baryon masses of the minor and major progenitors, respectively and M_{cold} is their total cold gas mass. The coefficient and index were originally

fixed to be consistent with the results of the Mihos & Hernquist (1996) simulations, but in HWT15 they are allowed to be free parameters. Although they are allowed to vary, the results are relatively consistent with the values assumed previously.

1.3.4 Dust Model

Star forming galaxies are known to be embedded in a rich dusty environment where their emission in the optical/UV are absorbed and re-emitted at longer wavelengths, the far-infrared/submillimetre. As a result, the overall emission of star-forming galaxies is obscured and affected by dust extinction. The radiative transfer model and dust reprocessing are complex and direct approaches to modelling the obscured emission. However, they require the grain details of the dust created by star formation and geometry and thus associate with high uncertainties. Simpler methods can be presented by taking into account the observational relation between the optical depth of dust in the galaxy disks and studying the connection between the UV/optical and far-infrared/submillimetre emission.

The current semi-analytic model uses the same approach as De Lucia & Blaizot (2007) where dust extinction is treated separately for a diffuse inter-stellar medium (ISM) component (following the work of Devriendt, Guiderdoni & Sadat (1999)) and for molecular clouds where stars form (following Charlot & Fall (2000)). The optical depth of the dust in each component is determined followed by some assumptions about its geometry, in order to compute the total extinction of the corresponding population.

In order to estimate the extinction of flux due to the ISM, the optical depth of diffuse dust in the galactic disk as a function of wavelength is given by:

$$\tau_{\lambda}^{ISM} = \left(\frac{A_{\lambda}}{A_{\nu}}\right)_{Z_{\odot}} (1+z)^{-1} \left(\frac{Z_{\rm gas}}{Z_{\odot}}\right)^{s} \times \left(\frac{< N_H >}{2.1 \times 10^{21}\,\mathrm{atoms\,cm^{-2}}}\right), \tag{1.9}$$

where $< N_H >$ shows the mean column density of hydrogen and is represented by :

$$< N_H >= \frac{M_{\rm cold}}{1.4\, m_{\rm p} \pi (a R_{\rm gas,d})^2} \mathrm{atoms\,cm^{-2}}, \tag{1.10}$$

where $R_{\rm gas,d}$ is the disk scale-length of the cold gas and $a = 1.68$ in order for $< N_H >$ to indicate the mass-weighted average column density of an exponential disk. Following the work of Guiderdoni & Rocca-Volmerange (1987) the extinction curve in Equation 1.9 is affected by the gas metallicity and using an interpolation from the Solar Neighbourhood and the Large and Small Magellanic Clouds, $s = 1.35$ when $\lambda < 2000$Å and $s = 1.6$ when $\lambda > 2000$Å. Equation 1.9 is more redshift dependent in comparison to the previous version of L-Galaxies such as $(1+z)^{-0.5}$ in Kitzbichler & White (2007) and $(1+z)^{-0.4}$

in Guo & White (2009). Motivated by observations (Steidel et al. (2004); Quadri et al. (2008)) and considering that relatively long-lived stars produce dust, this dependency means that for the same amount of cold gas and the same metal abundance, there is less dust at high redshift.

Another source of extinction only influences young stars and comes from the molecular clouds where they are formed. It assumes that this extinction exists for the stars that are younger than the lifetime of stellar birth clouds (taken to be 10^7 years). The relevant optical depth is given as:

$$\tau_\lambda^{BC} = \tau_\lambda^{ISM} \left(\frac{1}{\mu} - 1\right) \left(\frac{\lambda}{5000\mathring{A}}\right)^{-0.7}, \tag{1.11}$$

where μ has come from a random Gaussian distribution with standard deviation of 0.2 and a mean of 0.3, truncated at 0.1 and 1.

In order to calculate the final overall extinction from both components, every galaxy is assigned an inclination relative to the line of sight and a slab geometry is assumed for the disk associated with dust. Therefore, for each source the extinction is expressed as:

$$A_\lambda = -2.5 \log \left(\frac{1 - \exp^{-\tau_\lambda sec\theta}}{\tau_\lambda sec\theta}\right), \tag{1.12}$$

where θ is the angle of inclination of the galaxy correspond to the line of sight and τ_λ can be replaced by either $\tau\lambda^{ISM}$ or $\tau\lambda^{BC}$.

1.3.5 Stellar Population Synthesis Model and IMF

Stellar population synthesis models are a crucial part of galaxy formation theory as they link model prediction for stars (such as the masses, ages and metallicities) to the photometric properties. In this work I adopt Maraston (2005) stellar population synthesis model, however using Bruzual & Charlot (2003) leads to very similar results. The initial mass function (IMF) is one of the most important assumptions that should be included in SPS models. It holds information about the mass distribution of a newly-formed star population normalised by its total mass. SA model uses Chabrier initial function (Chabrier, 2003) that shows a better fit to observational data.

1.4 Infrared Continuum

The cosmic infrared background (CIB) is the total observed infrared emission from the galaxy formation era and forms around half of the extragalactic background light (EBL).

It indicates that a large fraction of radiation released by galaxies is absorbed by interstellar dust and re-emitted in the IR wavelengths, peaking at 150 micron.

Fig. 1.2 presents the EBL as a function of wavelength, showing that the cosmic infrared background (CIB) contributes roughly the same as the cosmic optical background. Since most of the interstellar dust is heated due to the radiation produced by young stars, the IR emission provides a powerful tracer of star formation activity over the history of the Universe. Although star formation rate (SFR) and infrared emission are strongly correlated, the calibration between these two quantities varies for all types of galaxies and depends on the optical depth of the dust surrounding a galaxy and the age of the stellar population. The efficiency of this expectation, indeed, depends on the contribution of young stars heating the dust and on the optical depth of the dust. The simplest physical assumption takes place when the radiation field in the UV-optical is dominated by young stars and the dust opacity is high everywhere. In this case, the FIR luminosity determines the bolometric luminosity of the source. Re-processing light from the young stars is not the only source contributing to the infrared background. The dust being heated by the radiation from the active galactic nucleus (AGN), however, based on multi-wavelength, especially X-ray studies, the AGN contribution appears to be insignificant, only $\sim$10 percent of the total (e.g. Almaini, Lawrence & Boyle (1999); Fardal et al. (2007)).

1.5 Star Formation History

Measuring the cosmic star formation history (SFH) is a fundamental key to understand galaxy formation and evolution and answer our big questions of when, where and in what conditions did stars form throughout cosmic history? The earlier studies were started on the rest-frame UV emission from galaxies as the SFR indicator (Lilly (1996); Madau et al. (1996); Steidel et al. (1999)) by using optically selected samples of galaxies at different redshifts. As discussed before, measuring SFR without taking into account the effect of dust attenuation is incomplete. Implementing the relations between UV extinction and UV spectral slope obtained from local objects (Calzetti et al. (1995); Meurer et al. (1995)) on the selected samples of galaxies at high redshift indicated that these galaxies should be affected by UV dust extinction; and the SFRs, derived from their rest-frame UV luminosities, should be corrected for dust extinction (Meurer, Heckman & Calzetti (1999); Steidel et al. (1999)).

There are 2 observational approaches to measuring the SFH of galaxies. The simpler one looks at the instantaneous SFR at different redshifts. In this approach SFR is

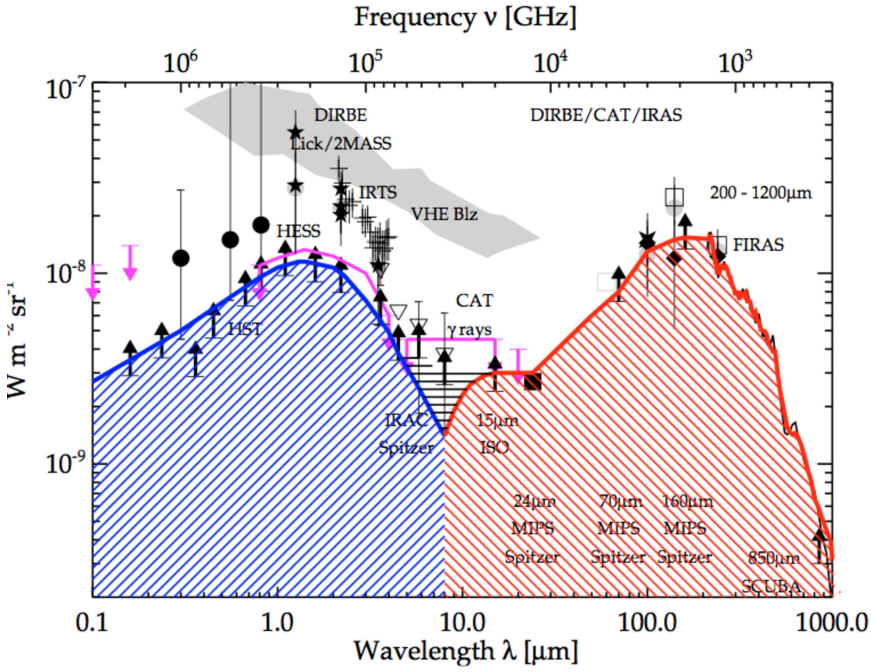

Figure 1.2: The extragalactic background light, showing the energy density as a function of wavelength as presented by Dole et al. (2006).

mostly obtained from UV luminosities (for an overview see, e.g. Kennicutt (1998); Calzetti (1999)).

The second approach can extrapolate the history from the fossil record of present-day galaxies and it is achieved by using the full available spectrum (Glazebrook et al. (2003); Panter, Heavens & Jimenez (2003); Heavens et al. (2004); Mathis, Charlot & Brinchmann (2006); Ocvirk et al. (2006); Cid Fernandes et al. (2007); Panter et al. (2007); Tojeiro et al. (2007)).

This thesis uses the archaeological approach, which is the latter method. The L-Galaxies semi-analytic model (hereafter simply L-Galaxies) is extended to keep a record of the SFHs of individual galaxies, with a bin-size that decreases with the look-back time. The introduction of SFHs into L-Galaxies allows comparison to observations, as well as reconstruction of galaxy magnitudes using arbitrary stellar population synthesis and dust models, in any band, in post-processing (Shamshiri et al., 2015).

In principle, the star formation history of a galaxy, its chemical evolution and its current dust content can be completely retrieved by using high enough quality observational data, proper modelling of the spectral energy distribution of stellar populations and dust extinction, followed by appropriate parameterisation. Several algorithms have been developed over the last decade to accomplish the above task: (see e.g. MOPED by Heavens et al. (2004); Panter et al. (2007); STECMAP by Ocvirk et al. (2006); STARLIGHT by Cid Fernandes et al. (2007) or ULYSS by Koleva et al. (2009)). In this thesis, I focus on the results obtained by VESPA (Tojeiro et al., 2007), a full spectral fitting code, that was applied to over 800,000 Sloan Digital Sky Survey DR7 galaxies (Abazajian et al., 2009) and resulted in the measurement of individual SFHs, metallicity histories and dust content (Tojeiro et al. (2007, hereafter TWH09). VESPA provides a robust comparison to our model predictions, due to the wide range of galaxies in SDSS DR7 and the time resolution of the recovered SFHs.

1.6 Herschel Space Observatory

The Herschel Space Observatory, (Pilbratt et al., 2010) was launched in 2009 before the supply of liquid helium coolant ran out in 2013, giving almost four years of glorious new data. In order to remove the infrared light that an Earth-orbiting telescope would be exposed to, Herschel was deployed at Lagrangian Point 2 (L2) a gravitationally stable point in the Earth-Sun system in the shadow of the Earth. Due to having the largest mirror in space at 3.5m wide and the focal plane cooled to the temperature of 1.4K, Herschel was

able to produce high resolution maps in the mid/far-infrared along with confusion limited maps in the far-infrared. The observatory covered the sub-millimetre and far infrared wavelength (55-672 μm) using three instruments, PACS (Photodetector Array Camera and Spectrometer,(Poglitsch et al., 2010)), SPIRE (Spectral and Photometric Imaging Receiver, (Griffin et al., 2010)) and a high resolution spectrometer, HIFI (Heterodyne Instrument for the Far-Infrared, (de Graauw et al., 2010)). PACS and SPIRE included both imaging cameras and medium resolution spectrometers while HIFI only consisted of a high resolution spectrometer.

This work uses the data and the results from the SPIRE photometry instrumentation. SPIRE covers the far-infrared wavelengths with photometric filters centred on 250, 350 and 500 μm and the Fourier-Transform Spectrometer (FTS) optimised for 200- 400μm (Valtchanov et al., 2014). Standard pixel sizes of an image from the photometer are of 6, 10 and 14 arcsec with PSF FWHMs of 18.15, 25.15 and 36.3 arcsec (Griffin et al., 2010). The receiving light is directed by eight different mirrors onto the detectors and filters that are used to construct the band pass (Ade et al., 2006).

1.7 Number counts

The statistical properties of galaxy populations are another important foundation for understanding galaxy formation and evolution. The number counts (i.e. the number density of galaxies as a function of flux) is the most basic statistic of galaxy populations. The number counts at far-infrared and sub-mm wavelengths display strong evolution, e.g. from ISO (Oliver et al. (2002); Héraudeau et al. (2004), and references therein), Spitzer (Shupe et al. (2008); Frayer et al. (2009), and references therein) and ground-based sub-mm surveys (Maloney et al. (2005); Scott et al. (2010), and references therein). The results from Herschel (Pilbratt et al., 2010) have offered a successful number count analysis to test and eliminate some existing models and provide a criterion upon which future models can be assessed. The data from Herschel suffers from confusion (Nguyen et al., 2010). Galaxies that are resolvable in the optical (SDSS) and near-infrared (Spitzer) may becoming unresolved point sources in the Herschel maps that are heavily blended together. This confusion means that in order to estimate number counts, caution has to be taken and the maps need to be corrected for flux boosting and incompleteness. Fig. 1.3 presents the number count calculated by Oliver et al. (2010) from counting galaxies directly after flux-deboosting and completeness corrections have been made. They compare the number counts with eight models including, pre-Spitzer (Xu et al., 2003), two based

on the ISO, SCUBA and Spitzer first results (Lagache et al., 2004) and 5 being more constrained by deep Spitzer, AzTEC, SCUBA and BLAST observations (Valiante et al. (2009); Le Borgne et al. (2009)). They have detected a sharp rise in the Euclidean normalised counts <100 mJy and directly resolved ∼15% of the infrared extra-galactic background (CIB) at 250 micron due to the source confusion of the Herschel/SPIRE instrument. As a result, the behaviour of the counts at faint flux density cannot be directly seen and thus, it is necessary to use statistical tools like $P(D)$ analysis (Condon (1974); Patanchon et al. (2009)) or stacking (Dole et al. (2006); Marsden et al. (2009)). Empirical models such as Béthermin et al. (2010) have used a stacking analysis in the SPIRE bands to determine deep counts per redshift slice combining the Herschel Multi-tiered Extragalactic Survey (HerMES) data (Oliver et al., 2012). The data mentioned above is used in this thesis to constrain our model prediction.

1.8 Luminosity Function

One of the most fundamental tools to probe distribution of galaxies over cosmological time is the luminosity function (LF). The LF describes the relative number of sources of different luminosities counted in characteristic volumes of the Universe. Measuring the LF over a range of redshifts is the most important technique for investigating the evolution of a galaxy population. When computed for different samples of galaxies, the LF can provide an essential contrast between the distributions of galaxies at different redshifts, in different environments or selected at different wavelengths. It is well known now that we cannot understand galaxy evolution without considering the energy absorbed by dust and re-radiated in the infrared or sub-millimetre (i.e. Genzel & Cesarsky (2000)). Since SF takes place within dusty environments, studying infrared and sub-millimetre emission is crucial to understand the full picture of star formation history over the cosmic time. With Herschel, a large number of high-z sources at the peak of their IR SED can be detected, however, source confusion and incompleteness limit our knowledge of the far-IR LF in the distant Universe.

Many authors have studied the IR LF using empirical and phenomenological models (e.g. Baugh et al. (2005); Fontanot et al. (2007); Lacey et al. (2011); Béthermin et al. (2011)). Empirical models provide a reasonable match to the data points, however, they are not a physical model and they do not increase our knowledge of the fundamental physics underpinning the galaxy formation. In contrast, SA models are powerful methods that describe the external and intrinsic properties of galaxies in physical details.

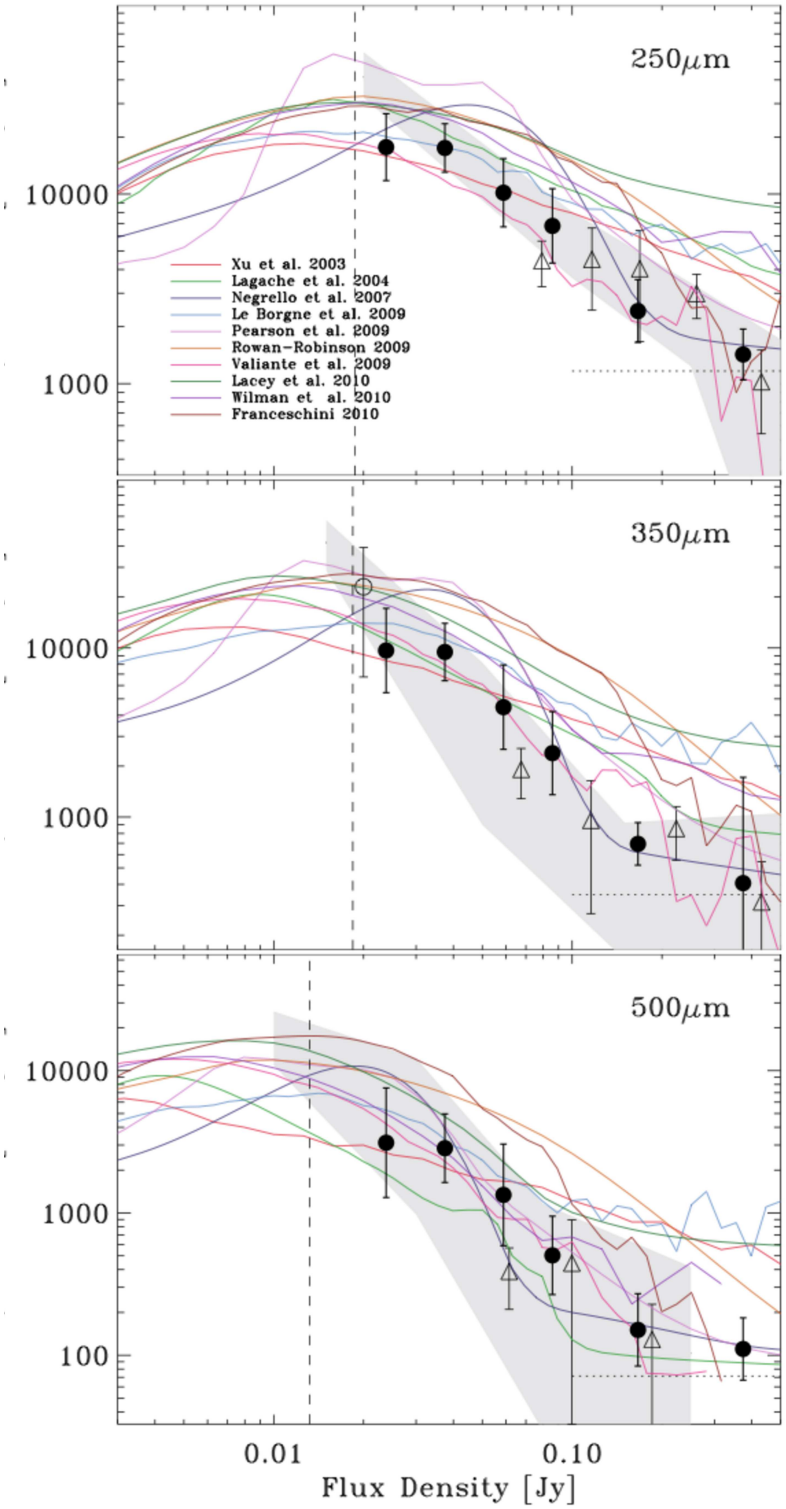

Figure 1.3: Number counts obtained from HerMES source catalogues. Filled circles are the mean number counts averaged over the following fields. GOODS-N & Lockman-North (faintest five bins only) and FLS& Lockman-SWIRE (brightest six bins only) after correction for flux-deboosting, completeness corrections and field-field error bars. Model fit to fluctuations of BLAST maps (omitting upper-limits,Patanchon et al. (2009)) - shaded region. Open triangles and open circle are from BLAST resolved counts (Béthermin et al., 2010) and (Khan et al., 2007) respectively. The figure is taken from Oliver et al. (2010).

Whilst semi analytical models generally fit observations well with a wide range of galaxy properties in the local Universe, there are some statistical quantities that are hard to properly explain. For example, these models mostly underestimate the number of sub-mm galaxies at high redshift and fail to match the bright end of the luminosity function (Cole et al. (2000); Baugh et al. (2005)). It is suggested that although model predictions of global properties such as the star formation rate and the total co-moving stellar mass are reasonably well matched with observations as a function of redshift, it is obvious that the physical mechanisms for star formation and evolution are not yet fully understood. The most recent version of the Durham semi-analytical model, GALFORM, assumes a top-heavy IMF for the stars formed in bursts, first proposed by Baugh et al. (2005), in order to solve the short-comings of number counts of galaxies and boosting the luminosity of galaxies in the sub-mm wavelength which broadly fit the observations (i.e. Lacey et al. (2011); Lacey et al. (2015)). The results of these models depend not only on invoking a top-heavy IMF but also detailed modelling of the absorption and the emission of radiation by dust.

This thesis attempts to predict IR luminosity and calculate the number density of infrared galaxies by proposing a simple and efficient model in the current version of L-Galaxies. Knowing that star formation rate (SFR) and infrared emission are strongly correlated, we adopt the simplest physical assumption that UV/optical radiation is dominated by young stars and the dust opacity is high everywhere which means that the FIR luminosity resolves the bolometric luminosity of the source (See the rest of the thesis for more details). The main benefit of our approach is avoiding the uncertainty due to the complexity of calculating the dust absorption using radiative transfer model.

As discussed above, in this thesis, I compute infrared luminosity function (IR LF) from a simulated SFR using the relation from Kennicutt (1998) (See chapter 3 the thesis for more details).

1.9 Overview of the papers

Previously, I described the importance of studying the SFR in galaxy evolution and gave the overall state of galaxy formation and evolution models related to infrared emission. In this thesis, I try to uncover the processes of galaxy evolution using the observable measurements, SFH, luminosity function and number counts so we will be able to establish an accurate physical prediction to analyse and understand the essential physical process behind galaxy evolution. I study and work mainly with the predictions of the Munich Semi

analytic model of SFR. I use the model outputs to calculate other observational properties in order to identify where L-Galaxies had been incomplete so far. I also contrast my result with available data in order to indicate whether the model is struggling to reproduce the data or if the observational results need reviewing. In the papers illustrated in this thesis, I estimate the SFH, number count and LF of galaxies and extend them to higher redshifts, where there is a lack of observational data.

The study of star formation rate as a function of cosmic time of the galaxies is the first fundamental stage for understanding the astrophysics behind the formation and evolution of galaxies. The observationally-measured SFHs are vastly affected by the quality of the stellar spectra and are sensitive to the input stellar population synthesis models. Therefore, our prediction in principle is a good constraint and benchmark for the credibility of their results.

Considering the above, Paper I, concentrates on the star formation history from the fossil record of current-day galaxies. This work is adopted to the Munich SAM, known as L-Galaxies, and is also the third in the series of articles aiming to introduce galaxy formation in the Plank cosmology. Our result is reasonably consistent with the SDSS-7 catalogue. The simulated and SDSS galaxies show that the SFH gets steeper with increasing stellar mass. In other words, more massive galaxies form their stars earlier and have lower current SSFRs than lower-mass galaxies. We model how the star formation rate of galaxies evolves with redshift and also show the clear evidence of down-sizing in that more massive galaxies become quiescent first.

In contrast, the SFHs of blue and red galaxies show poor consistency with the data and simulations. The SA models show much more drastic deviations between the two populations than the observed galaxies from the VESPA database. In fact, the observational SFHs for red and blue galaxies begin to deviate from the youngest stars of age less than 0.3 Gyr, whereas the first model begins to differ from the stars of age about 5 Gyr.

I have estimated the model SFHs using different constraints with respect to the colour and mass and made all the comparisons with Vespa algorithm as applied to the SDSS-7 catalogue. Many discussions and suggestions were accomplished by both my supervisors and other collaborators in different meetings with them. Dr. Bruno Henriques was in charge of testing and validation of the post-processing method and compared the magnitude driven by full resolution and binned SFHs and as a complement, Prof. Peter Thomas made the predictions from our models of how the star formation rate of galaxies evolves with redshift.

In the first paper, I did all the study and analysis of the model. I provided all the comparison figures between the L-Galaxies and SDSS sources along with interpretation under the supervision of Prof. Peter Thomas. This paper is already accepted by the MNRAS journal.

The second paper started as a continuation of testing the L-Galaxies prediction and its capability as a major constraining model, which was launched in my pervious work (Shamshiri et al., 2015). Since starlight is absorbed by dust, reprocessed and emitted in the rest-frame mid/far-IR, studying mid/far-IR emission provides a robust constraint on the total (or intrinsic) star-formation activity. Therefore, the reliability of the dust model (which includes the formation and demolition of dust, and its effect on the intrinsic spectral energy distribution) is crucial in order for direct comparison of the predictions from galaxy formation models with the observed IR number count and luminosity function. In the second paper, I used a simple and effective model to compute IR luminosity from obscured star formation. We converted the simulated SFR to luminosity density as described in Kennicutt (1998) then employed the SED library (as described in Béthermin et al. (2012)) to normalise IR LF to monochromatic flux using the equation expressed in Section 3.3. Thus, the result is used to produce the mock catalogue, which now contains fluxes associated with the SFR in the catalogue. As a result L-Galaxies now is extended to predict IR part of spectrum and able to provide statistical analysis over the infrared wavelengths.

Comparing the number counts density to observation indicates that my model prediction of number count lies below the observation data points (e.g Béthermin et al. (2012) and Oliver et al. (2010)). L-Galaxies is also able to provide the infrared luminosities since they are essential for understanding galaxy formation and evolution, and is a complementary stage for interpreting our predictions of number counts. I then calculated and characterise the IR LF. The simulated and observed LFs present an overall match at lower redshift, however, the simulated one is left behind by increasing redshift.

The luminosity function is fitted to the data points in order to have a better understanding of what is needed to improve the model and how number counts are affected. The condition I used for my fitting operation was that the total flux density remains constant pre and post fitting. For fitting, I also adopted the physical approach when the total flux density remains constant before and after fitting. The result clearly fits better for all the redshifts although the simulation presents a higher peak compared to the observations. These fitted parameters are applied later on to mono-

chromatic fluxes, which lead to boosting the fluxes but not necessarily matching the data points. I discuss all the possible results that can cause this exceed of flux density in the second paper. I produced all the figures of the paper my self. Prof. Peter Thomas and Prof. Seb Oliver have collaborated by giving important comments and discussions.

Bibliography

Abazajian K. N. et al., 2009, ApJS, 182, 543 Cited on 12

Ade P. A. R., Pisano G., Tucker C., Weaver S., 2006, in Proc. SPIE, Vol. 6275, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, p. 62750U Cited on 13

Almaini O., Lawrence A., Boyle B. J., 1999, MNRAS, 305, L59 Cited on 10

Angulo R. E., White S. D. M., 2010, MNRAS, 405, 143 Cited on 4

Baugh C. M., 2006, Reports on Progress in Physics, 69, 3101 Cited on 1

Baugh C. M., Lacey C. G., Frenk C. S., Granato G. L., Silva L., Bressan A., Benson A. J., Cole S., 2005, MNRAS, 356, 1191 Cited on 14, 16

Béthermin M., Dole H., Cousin M., Bavouzet N., 2010, A&A, 516, A43 Cited on 14, 15

Béthermin M., Dole H., Lagache G., Le Borgne D., Penin A., 2011, A&A, 529, A4 Cited on 14

Béthermin M. et al., 2012, A&A, 542, A58 Cited on 18

Bigiel F., Leroy A., Walter F., Brinks E., de Blok W. J. G., Madore B., Thornley M. D., 2008, AJ, 136, 2846 Cited on 7

Boylan-Kolchin M., Springel V., White S. D. M., Jenkins A., Lemson G., 2009, MNRAS, 398, 1150 Cited on 4

Bruzual G., Charlot S., 2003, MNRAS, 344, 1000 Cited on 9

Calzetti D., 1999, Ap&SS, 266, 243 Cited on 12

Calzetti D., Bohlin R. C., Kinney A. L., Storchi-Bergmann T., Heckman T. M., 1995, ApJ, 443, 136 Cited on 10

Chabrier G., 2003, PASP, 115, 763 Cited on 7, 9

Charlot S., Fall S. M., 2000, ApJ, 539, 718 Cited on 8

Cid Fernandes R., Asari N. V., Sodré L., Stasińska G., Mateus A., Torres-Papaqui J. P., Schoenell W., 2007, MNRAS, 375, L16 Cited on 12

Cole S., Lacey C. G., Baugh C. M., Frenk C. S., 2000, MNRAS, 319, 168 Cited on 16

Condon J. J., 1974, ApJ, 188, 279 Cited on 14

de Graauw T. et al., 2010, A&A, 518, L6 Cited on 13

De Lucia G., Blaizot J., 2007, MNRAS, 375, 2 Cited on 8

De Lucia G., Kauffmann G., White S. D. M., 2004, MNRAS, 349, 1101 Cited on 6

Devriendt J. E. G., Guiderdoni B., Sadat R., 1999, A&A, 350, 381 Cited on 8

Dole H. et al., 2006, A&A, 451, 417 Cited on 11, 14

Efstathiou G., 2002, MNRAS, 332, 193 Cited on 1

Fardal M. A., Katz N., Weinberg D. H., Davé R., 2007, MNRAS, 379, 985 Cited on 10

Fontanot F., Monaco P., Silva L., Grazian A., 2007, MNRAS, 382, 903 Cited on 14

Frayer D. T. et al., 2009, AJ, 138, 1261 Cited on 13

Frieman J. A., Turner M. S., Huterer D., 2008, ARA&A, 46, 385 Cited on 1

Genzel R., Cesarsky C. J., 2000, ARA&A, 38, 761 Cited on 14

Glazebrook K. et al., 2003, ApJ, 587, 55 Cited on 12

Griffin M. J. et al., 2010, A&A, 518, L3 Cited on 13

Guiderdoni B., Rocca-Volmerange B., 1987, A&A, 186, 1 Cited on 8

Guo Q. et al., 2011, MNRAS, 413, 101 Cited on 7

Guo Q., White S. D. M., 2009, MNRAS, 396, 39 Cited on 9

Hanany S. et al., 2000, ApJ, 545, L5 Cited on 1, 2

Heavens A., Panter B., Jimenez R., Dunlop J., 2004, Nature, 428, 625 Cited on 12

Henriques B., White S., Thomas P., Angulo R., Guo Q., Lemson G., Springel V., Overzier R., 2015, MNRAS, 451, 2263 Cited on 4, 5

Héraudeau P. et al., 2004, MNRAS, 354, 924 Cited on 13

Hinshaw G. et al., 2007, ApJS, 170, 288 Cited on 2

Jones W. C. et al., 2006, ApJ, 647, 823 Cited on 1, 2

Kauffmann G., 1996, MNRAS, 281, 487 Cited on 7

Kauffmann G., Colberg J. M., Diaferio A., White S. D. M., 1999, MNRAS, 303, 188 Cited on 7

Kennicutt, Jr. R. C., 1998, ARA&A, 36, 189 Cited on 12, 16, 18

Khan S. A. et al., 2007, ApJ, 665, 973 Cited on 15

Kitzbichler M. G., White S. D. M., 2007, MNRAS, 376, 2 Cited on 8

Koleva M., Prugniel P., Bouchard A., Wu Y., 2009, A&A, 501, 1269 Cited on 12

Lacey C. G., Baugh C. M., Frenk C. S., Benson A. J., 2011, MNRAS, 412, 1828 Cited on 14, 16

Lacey C. G. et al., 2015, ArXiv e-prints Cited on 16

Lagache G. et al., 2004, ApJS, 154, 112 Cited on 14

Le Borgne D., Elbaz D., Ocvirk P., Pichon C., 2009, A&A, 504, 727 Cited on 14

Leroy A. K., Walter F., Brinks E., Bigiel F., de Blok W. J. G., Madore B., Thornley M. D., 2008, AJ, 136, 2782 Cited on 7

Lilly S., 1996, in APS Meeting Abstracts Cited on 10

Madau P., Ferguson H. C., Dickinson M. E., Giavalisco M., Steidel C. C., Fruchter A., 1996, MNRAS, 283, 1388 Cited on 10

Maloney P. R. et al., 2005, ApJ, 635, 1044 Cited on 13

Maraston C., 2005, MNRAS, 362, 799 Cited on 9

Marsden G. et al., 2009, ApJ, 707, 1729 Cited on 14

Mathis H., Charlot S., Brinchmann J., 2006, MNRAS, 365, 385 Cited on 12

Meurer G. R., Heckman T. M., Calzetti D., 1999, ApJ, 521, 64 Cited on 10

Meurer G. R., Heckman T. M., Leitherer C., Kinney A., Robert C., Garnett D. R., 1995, AJ, 110, 2665 Cited on 10

Mihos J. C., Hernquist L., 1996, ApJ, 464, 641 Cited on 8

Nguyen H. T. et al., 2010, A&A, 518, L5 Cited on 13

Ocvirk P., Pichon C., Lançon A., Thiébaut E., 2006, MNRAS, 365, 74 Cited on 12

Oliver S. et al., 2002, MNRAS, 332, 536 Cited on 13

Oliver S. J. et al., 2012, MNRAS, 424, 1614 Cited on 14

Oliver S. J. et al., 2010, A&A, 518, L21 Cited on 13, 15, 18

Panter B., Heavens A. F., Jimenez R., 2003, MNRAS, 343, 1145 Cited on 12

Panter B., Jimenez R., Heavens A. F., Charlot S., 2007, MNRAS, 378, 1550 Cited on 12

Patanchon G. et al., 2009 Cited on 14, 15

Percival W. J. et al., 2001, MNRAS, 327, 1297 Cited on 1

Pilbratt G. L. et al., 2010, A&A, 518, L1 Cited on 12, 13

Poglitsch A. et al., 2010, A&A, 518, L2 Cited on 13

Press W. H., Schechter P., 1974, ApJ, 187, 425 Cited on 4

Quadri R. F., Williams R. J., Lee K.-S., Franx M., van Dokkum P., Brammer G. B., 2008, ApJ, 685, L1 Cited on 9

Scott K. S. et al., 2010, MNRAS, 405, 2260 Cited on 13

Shamshiri S., Thomas P. A., Henriques B. M., Tojeiro R., Lemson G., Oliver S. J., Wilkins S., 2015 Cited on 12, 18

Shupe D. L. et al., 2008, AJ, 135, 1050 Cited on 13

Smoot G. F. et al., 1992, ApJ, 396, L1 Cited on 1

Somerville R. S., Primack J. R., Faber S. M., 2001, MNRAS, 320, 504 Cited on 7

Springel V. et al., 2005, Nature, 435, 629 Cited on 1, 2, 3, 4

Steidel C. C., Adelberger K. L., Giavalisco M., Dickinson M., Pettini M., 1999, ApJ, 519, 1 Cited on 10

Steidel C. C., Shapley A. E., Pettini M., Adelberger K. L., Erb D. K., Reddy N. A., Hunt M. P., 2004, ApJ, 604, 534 Cited on 9

Sutherland R. S., Dopita M. A., 1993, ApJS, 88, 253 Cited on 6

Tegmark M. et al., 2004, ApJ, 606, 702 Cited on 1

Tojeiro R., Heavens A. F., Jimenez R., Panter B., 2007, MNRAS, 381, 1252 Cited on 12

Valiante E., Lutz D., Sturm E., Genzel R., Chapin E. L., 2009, ApJ, 701, 1814 Cited on 14

Valtchanov I. et al., 2014, Experimental Astronomy, 37, 207 Cited on 13

White S. D. M., Frenk C. S., 1991, ApJ, 379, 52 Cited on 5

White S. D. M., Rees M. J., 1978, MNRAS, 183, 341 Cited on 2

Xu C. K., Lonsdale C. J., Shupe D. L., Franceschini A., Martin C., Schiminovich D., 2003, ApJ, 587, 90 Cited on 13

Chapter 2

Paper1

Galaxy formation in the Planck cosmology - II. Star formation histories and post-processing magnitude reconstruction

Sorour Shamshiri, Peter A. Thomas Bruno M. Henriques, Rita Tojeiro Gerard Lemson, Seb J. Oliver, Stephen Wilkins

2.1 Abstract

We adapt the L-Galaxies semi-analytic model to follow the star formation histories (SFH) of galaxies – by which we mean a record of the formation time and metallicities of the stars that are present in each galaxy at a given time. We use these to construct stellar spectra in post-processing, which offers large efficiency savings and allows user-defined spectral bands and dust models to be applied to data stored in the Millennium data repository.

We contrast model SFHs from the Millennium Simulation with observed ones from the VESPA algorithm as applied to the SDSS-7 catalogue. The overall agreement is good, with both simulated and SDSS galaxies showing a steeper SFH with increased stellar mass. The SFHs of blue and red galaxies, however, show poor agreement between data and simulations, which may indicate that the termination of star formation is too abrupt in the models.

The mean star-formation rate (SFR) of model galaxies is well-defined and is accurately modelled by a double power law at all redshifts: $\mathrm{SFR} \propto 1/(x^{-1.39} + x^{1.33})$, where $x = (t_\mathrm{a} - t)/3.0\,\mathrm{Gyr}$, t is the age of the stars and t_a is the lookback time to the onset of galaxy

formation; above a redshift of unity, this is well approximated by a gamma function: $\mathrm{SFR} \propto x^{1.5} e^{-x}$, where $x = (t_\mathrm{a} - t)/2.0\,\mathrm{Gyr}$. Individual galaxies, however, show a wide dispersion about this mean. When split by mass, the SFR peaks earlier for high-mass galaxies than for lower-mass ones, and we interpret this downsizing as a mass-dependence in the evolution of the quenched fraction: the SFHs of star-forming galaxies show only a weak mass dependence.

2.2 Introduction

Understanding the astrophysics behind the formation and evolution of galaxies is an important goal in modern astronomy. One of the most fundamental probes of that physics is the star formation rate as a function of cosmic time. In this paper we contrast predicted and observed star formation histories (hereafter, SFHs) of galaxies, and explore the expected range of SFHs at high redshift.

Two main observational approaches are used to infer the SFH of galaxies. One can look at the instantaneous star formation rate as a function of cosmic time (for an overview see, e.g. Kennicutt, 1998; Calzetti, 1999), or one can deduce the history from the fossil record of current-day galaxies. The two techniques in fact measure slightly different things, with the relationship between the two depending upon the merging history of the galaxies.

This paper focuses upon the archaeological approach. We have extended the L-Galaxies semi-analytic model (hereafter simply L-Galaxies) to keep a record of the SFHs of individual galaxies, with a bin-size that increases with increasing age for the stars. The resulting data have been made available as part of the public data release (DR) of the Millennium Simulation* (Lemson & The Virgo Consortium, 2006) that accompanies the latest implementation of L-Galaxies (Henriques et al., 2014).

We note that the SFH as defined in this paper (the history of star formation of all the stars that end up in the galaxy at some particular time) is related to, but distinct from, the variation in stellar mass of a galaxy over time. The latter follows only the history of the main component of the galaxy (along the "main branch" of the merger tree) and has been investigated for the Millennium Simulation by Cohn & van de Voort (2015). The difference between the two reflects the merger history of galaxies.

The term SFH is often loosely used in papers without being defined. Observationally, the only *direct* measure of SFHs corresponds to that described in this paper, i.e. the

*`http://gavo.mpa-garching.mpg.de/MyMillennium/`

distribution of formation times of all the stars that make up the galaxy.* Measures of the SFH of the main galactic component can only be inferred statistically by observing populations of galaxies at different redshifts and making some assumptions about merger rates as a function of stellar mass and environment. In principle, the former method is much cleaner, as it is free from these model assumptions; however, in practice, the inversion of the stellar spectra is highly degenerate and sensitive to the input stellar population synthesis models, and can lead to implausible results if some model constraints are not imposed.

As well as enabling comparison to observations, the introduction of SFHs into L-Galaxies allows reconstruction of galaxy magnitudes using arbitrary stellar population synthesis and dust models, in any band, in post-processing, and we investigate the accuracy of this approach. In addition, having SFHs is a prerequisite for a correct, time-resolved treatment of galactic chemical enrichment, as described in Yates et al. (2013).

The star formation history of a galaxy, its chemical evolution, and its current dust content, can in principle be fully recovered with high enough quality observational data, suitable modelling of the spectral energy distribution of stellar populations and dust extinction, and appropriate parametrization. Several algorithms have been developed over the last decade that attempt to do the above in the most robust way (see e.g. MOPED by Heavens et al. 2004; Panter et al. 2007, STECMAP by Ocvirk et al. 2006, STARLIGHT by Cid Fernandes et al. 2004, 2005 or ULYSS by Koleva et al. 2009). In this paper we focus on the results obtained by VESPA (Tojeiro et al., 2007), a full spectral fitting code that was applied to over 800,000 Sloan Digital Sky Survey DR7 galaxies (SDSS Collaboration, 2009). The resulting data base of individual SFHs, metallicity histories and dust content is publicly available and described in Tojeiro et al. (2009, hereafter TWH09). The wide range of galaxies in SDSS DR7 and the time resolution of the SFHs published in the data base make it ideal for a detailed comparison with our model predictions.

At higher redshift, observational data are scarce. Nevertheless, we make predictions from our models of how the star formation rate of galaxies evolves with redshift, and we show how downsizing arises from a mass-dependence of the rate at which galaxies are quenched.

In outline, the main aims of this paper are:

- to briefly overview the L-Galaxies and VESPA algorithms – Sections 3.3.1 and

*We note that this would better be described by the phrase "stellar age distribution" rather than star formation history, but the latter phrase predominates in the literature.

2.3.2;

- to describe the L-Galaxies SFH binning method – Section 2.3.3;
- to test how well post-processing can reconstruct magnitudes – Section 2.4;
- to compare model galaxies from L-Galaxies with results from the VESPA catalogue – Section 2.5;
- to investigate the variety of SFHs that we find in our model galaxies – Section 2.6;
- to provide a summary of our key results – Section 2.7.

2.3 Methods

2.3.1 L-Galaxies

In this work we use the latest version of the Munich semi-analytic (SA) model, L-Galaxies, as described in Henriques et al. (2014). This gives a good fit to, amongst other things, the observed evolution of the mass and luminosity functions of galaxies, the fraction of quenched galaxies, the star formation versus stellar mass relation (at least at $z < 2$), the Tully–Fisher relation, metallicities, black hole masses, etc. We refer to this standard model as HWT15. The improvements obtained in terms of the evolution of the abundance and red fractions of galaxies as a function of stellar mass in this model result from: a supernova feedback model in which ejected gas is allowed to fall back on to the galaxy on a time-scale that scales inversely with halo virial mass (Henriques et al., 2013); and a lower star formation threshold and weaker environmental effects both reducing the suppression of star formation in dwarf galaxies.

L-Galaxies, as with other SA models, follows the growth of galaxies within the framework of a merger tree of dark matter haloes. We construct this tree from the Millennium Simulation (Springel et al., 2005), scaled using the method of Angulo & White (2010) to the *Planck* cosmology (Planck Collaboration XVI, 2014): $\Omega_{\rm m} = 0.315$, $\Omega_\Lambda = 0.683$, $\Omega_{\rm b} = 0.0488$, $n_{\rm s} = 0.958$, $\sigma_8 = 0.826$, $h = 0.673$. This then gives a box size of $480.3\,h^{-1}$ Mpc and a particle mass of $9.61 \times 10^8 h^{-1}$ $\mathrm{M}_\odot$. The merger tree is constructed from 58 snapshots,* each of which is subdivided into 20 integration timesteps. The snapshots are unevenly spaced, such that the time resolution is higher at high redshift, but a typical timestep is of the order of $1\text{–}2\times10^7$ yr. We note that the SA model has been

*Five of the original *Millennium Simulation* snapshots lie at $z < 0$ after scaling.

implemented on both the higher resolution Millennium-II (Boylan-Kolchin et al., 2009) and larger volume Millennium-XXL simulations (Angulo et al., 2014) although we do not make use of those simulations in the current work.

Prior to HWT15, galaxy luminosities were determined by accumulating flux in different spectral bands throughout the time-evolution of each galaxy. In the new model those fluxes are recovered to high accuracy in post-processing, by simply recording the star formation and metallicity history in a relatively small number of time-bins. The new method is introduced following the work outlined and presented in the current paper. The DR that accompanies this series of papers records the SFHs, allowing the user flexibility to define their own bands and dust models.

2.3.2 VESPA

The spectrum of a galaxy, in the absence of dust, can be described as the linear superposition of the spectra of the stellar populations of different ages and metallicities that exist in the galaxy. The deconvolution of a galaxy's spectrum into a star formation and metallicity history is in principle trivial, but complicated by noisy or incomplete data and limitations in the modelling. Ocvirk et al. (2006) showed how the problem quickly becomes ill conditioned as noise increases in the data, and that the risk of overparametrizing a galaxy is high. VESPA takes into account the noise and data quality of each individual galaxy and uses an algebraic approach to estimate how many linearly independent components one can extract from each observed spectrum, thereby avoiding fitting the noise rather than the signal - see Tojeiro et al. (2007) for details.

VESPA recovers the SFH of a galaxy in 3 to 16 age bins (depending on the quality of the spectra), logarithmically spaced between 0.002 Gyr and the age of the Universe. For each age bin, VESPA returns the total mass formed within the bin and the mass-weighted metallicity of the bin, together with an estimate of the dust content of the galaxy. As we always compare our model predictions to the mean SFH of large ensembles of galaxies, we choose to use the fully-resolved SFHs published in the data base of TWH09. Whereas we expect these to be dominated by the noise on each individual galaxy, the mean over a large ensemble has been shown to be robust (Panter, Heavens & Jimenez, 2003). In this paper, we will compare to mean SFHs for different galaxy populations, as described in the text.

2.3.3 The SFH binning algorithm

As mentioned in Section 2.3.1, the Millennium merger trees are constructed from 58 snapshots, each of which is separated into 20 integration timesteps. To follow the history of star formation, we introduce extra arrays to carry information on the mass and metallicity of stars in each component of the galaxy (disc, bulge, intracluster mass) as a function of cosmic time. To save the information over all 1160 timesteps would consume too much memory and is unnecessary. Instead, we wish to use a high resolution for the recent past (when the stellar population is rapidly evolving) and a lower one at more distant times. To do so, we adopt the following procedure (see Fig. 2.1).

Starting at high redshift, on each timestep a new bin is created to hold the SFH information. Whenever the number of bins of a particular resolution exceeds $N_{\rm max}$ (where in the diagram, for the purposes of illustration, $N_{\rm max} = 2$), then the two oldest bins are merged together to form a new bin of twice the size – this may result in a cascade of mergers at successively higher levels (in the figure, these mergers are represented by red columns joined by braces to the merger product). In this way, the number of bins at each size grows from 1 to $N_{\rm max}$, then oscillates between $N_{\rm max}$ and $N_{\rm max}-1$. The total number of bins required does not exceed the smallest integer greater than $N_{\rm max}\log_2(N_{\rm step}/N_{\rm max}+1)$, where $N_{\rm step}$ is the number of timesteps. Table 2.1 shows, for the *Millennium Simulation*, using 20 steps within each of 58 snapshots, the maximum number of bins required, and, at $z = 0$, the actual number of SFH bins and their minimum and maximum size in years.* Note that all choices of $N_{\rm max} \geq 2$ have the same minimum bin-size, equal to that of the original timesteps; what differs is the number of bins that are resolved at that highest resolution.

We have investigated the sensitivity of our results to the number of bins and conclude that $N_{\rm max} = 2$ gives the best balance between data-size and accuracy: that is the value used in the Millennium data base and, unless mentioned otherwise, in the results presented below.

2.4 Post-processing of magnitudes

In most SA models, and in L-Galaxies prior to this work, galaxy luminosities are computed by adding the flux in different bands throughout the time-evolution of each galaxy. This calculation generally requires interpolating between values in large stellar population

*These numbers are unchanged for the 63 snapshots required to use the *Millennium Simulation* with the original *WMAP*-1 cosmology.

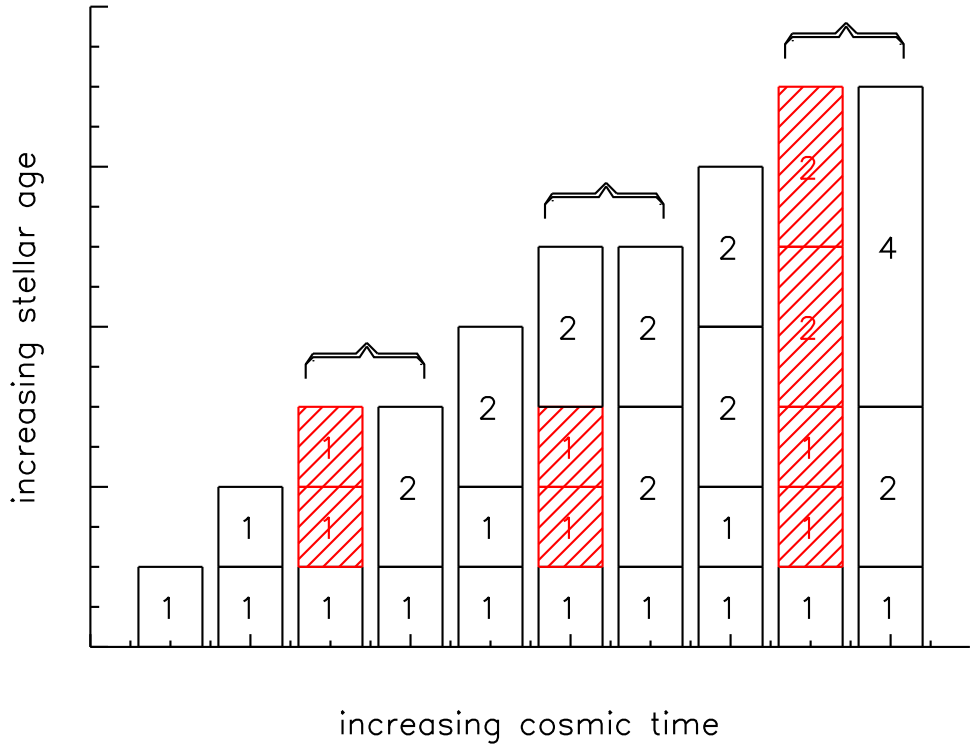

Figure 2.1: The evolution of the SFH bins for $N_{\rm max} = 2$. The x-axis represents successive timesteps, starting from high redshift and moving towards the present day. The y-axis represents bins of stellar age, starting with stars that are newly-created at that time and looking back into the past. The numbers within each bin represent the number of timesteps that have been merged together to produce that bin. The columns that are bracketed together show different arrangements of the data at a single cosmic time. The shaded, red stacks represent transient structures in which some bins merge together to produce the black stacks to their right.

Table 2.1: For the *Millennium Simulation*, using 1160 timesteps, this table shows, for different choices of $N_{\rm max}$: $N_{\rm tot}$ – the maximum number of SFH bins required; $N_{z=0}$ – the number of populated bins at $z = 0$; $\Delta t_{\rm min}/{\rm yr}$ – the minimum bin-size in years at $z = 0$; $\Delta t_{\rm max}/{\rm yr}$ – the maximum bin-size in years at $z = 0$.

$N_{\rm max}$	$N_{\rm tot}$	$N_{z=0}$	$\Delta t_{\rm min}/{\rm yr}$	$\Delta t_{\rm max}/{\rm yr}$
1	11	7	6.0×10^7	1.1×10^{10}
2	19	16	1.5×10^7	2.1×10^9
3	27	23	1.5×10^7	1.6×10^9
4	34	31	1.5×10^7	5.6×10^8

Table 2.2: This table shows the root-mean-square difference between magnitudes calculated on-the-fly during the running of the code and those calculated in post-processing.

$z = 0$	Without dust			With dust		
	$N_{\max} = 1$	$N_{\max} = 2$	$N_{\max} = 4$	$N_{\max} = 1$	$N_{\max} = 2$	$N_{\max} = 4$
GALEX FUV	3.02	0.29	0.16	2.53	0.38	0.29
GALEX NUV	1.79	0.05	0.03	1.47	0.16	0.16
SDSS u	0.76	0.03	0.01	0.66	0.06	0.05
SDSS g	0.50	0.02	0.00	0.46	0.03	0.02
SDSS z	0.36	0.02	0.00	0.33	0.02	0.02
VISTA J	0.35	0.02	0.00	0.32	0.02	0.02
VISTA K_s	0.33	0.01	0.01	0.31	0.02	0.02
IRAC 3.6$_{\mu\mathrm{m}}$	0.38	0.01	0.00	0.36	0.03	0.02
$z = 2$	without dust			with dust		
	$N_{\max} = 1$	$N_{\max} = 2$	$N_{\max} = 4$	$N_{\max} = 1$	$N_{\max} = 2$	$N_{\max} = 4$
GALEX FUV	1.74	0.04	0.01	1.14	0.27	0.26
GALEX NUV	1.39	0.02	0.00	0.95	0.21	0.20
SDSS u	0.88	0.02	0.00	0.63	0.13	0.12
SDSS g	0.64	0.01	0.00	0.47	0.09	0.09
SDSS z	0.39	0.01	0.00	0.33	0.04	0.04
VISTA J	0.33	0.01	0.00	0.30	0.02	0.02
VISTA K_s	0.32	0.01	0.00	0.30	0.01	0.01
IRAC 3.6μm	0.31	0.01	0.00	0.28	0.02	0.02

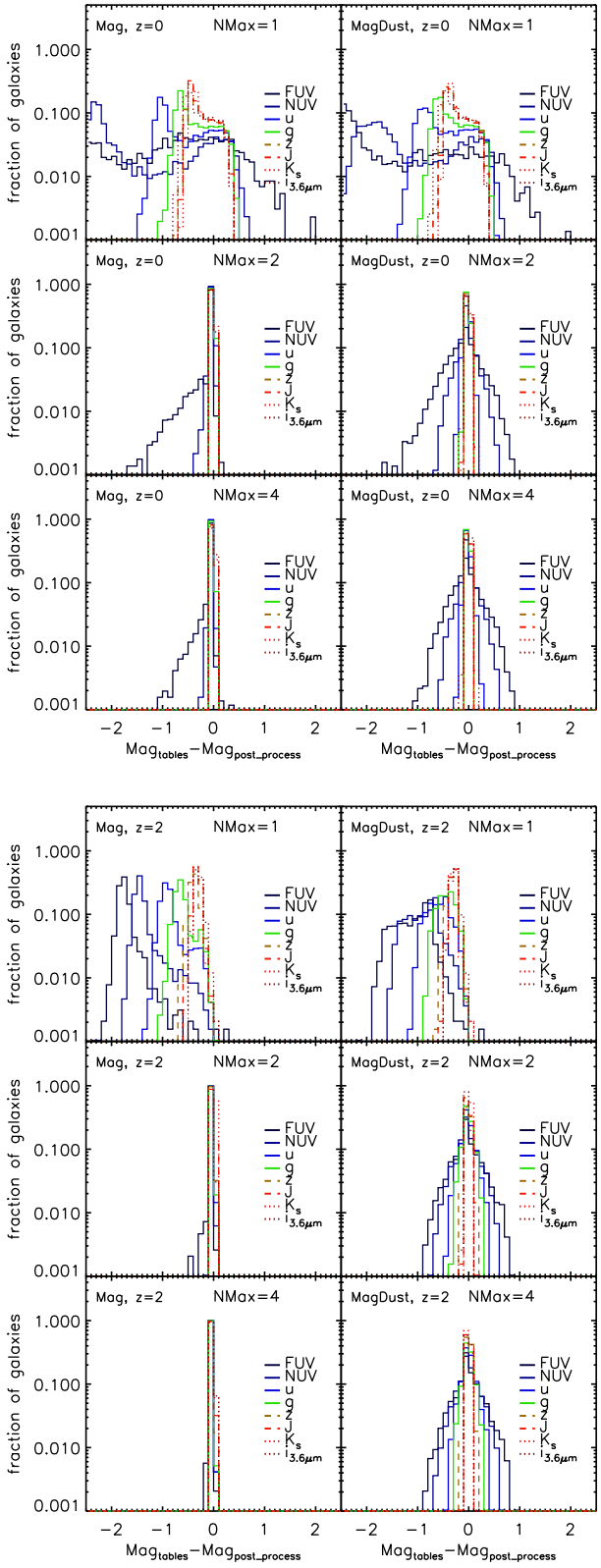

Figure 2.2: The difference between photometric properties calculated using full resolution and binned SFHs at $z = 0$ (top) and $z = 2$ (bottom). The left-hand column in each set of panels correspond to intrinsic magnitudes while the right takes into account dust extinction. From top to bottom the resolution of the binning is increased: $N_{\max} = 1$ (top), 2 (middle) and 4 (bottom).

synthesis tables and represents a large fraction, and in some cases the majority, of the computational time for the entire galaxy formation model. The problem is aggravated as different types of magnitudes (dust corrected, observer-frame) for additional components (e.g. disc, bulge, intra-cluster light) are included.

These difficulties can, in principle, be circumvented by storing the star formation and metallicity histories for different components of the galaxies, and using them to compute emission in post-processing. Ideally, this history would be stored for all the intermediate steps between output snapshots for which galaxy properties are computed. However, memory constraints make this infeasible. For our current set up, for example, it would require storage of up to 2320 values for each galaxy component (58 snapshots, 20 intermediate steps per snapshot, star formation and metallicity).

An alternative, tested in this paper, is to store the histories in bins that grow in size for older populations, as described in Section 2.3.3. Since the emission properties of populations vary on significantly longer time scales for old populations this can in principle allow us to maintain accuracy. To validate the method we compare the theoretical emission from galaxies computed both with full resolution on-the-fly and from star formation and metallicity history bins that are merged together for older populations. We assume that star formation occurs at a time corresponding to the mid-point of each SFH bin. To spread star formation out over the time-bin would be equivalent to using a larger number of timesteps (which we have also tested) and makes little difference except in the UV.

Fig. 2.2 shows the difference between photometric properties calculated using full resolution and binned SFHs for $z = 0$ and $z = 2$. In both sets of panels the left-hand column corresponds to intrinsic magnitudes while the right-hand column takes into account dust extinction. From top to bottom the resolution of the binning is increased: $N_{\rm max} = 1$, 2 and 4. While large differences between the two methods are seen for the lowest resolution, the figures show that good convergence is achieved for $N_{\rm max} = 2$ or more.

Table 2.2 shows the root-mean-square (rms) difference between magnitudes calculated on-the-fly during the running of the code (i.e. using the finest possible time resolution) and those calculated in post-processing, using different numbers of SFH bins. Quantitatively, for $N_{\rm max} = 2$, at $z = 0$, the rms difference between the two methods is less than 0.05 for all intrinsic magnitudes except the far-UV, for which it is approximately 0.29.* At $z = 2$, the mean difference is less than 0.04 in all bands.

The increased accuracy in the far-UV at high redshift results from the higher accuracy

*These values are reduced to 0.03 and 0.16, respectively, when $N_{\rm max} = 4$ is used.

in post-processing at a time in which fewer bins were merged. Emission in this part of the spectrum is dominated by extremely young populations for which even a slightly different formation time results in a large variation in predicted flux. The two methods thus differ in detail, but have the same statistical properties. If one is interested in the precise UV flux of a particular galaxy then that can be recovered by using finer time-bins. We have checked that, keeping $N_{\rm max} = 2$ but using a finer timestep (which adds very few SFH bins) improves the agreement between the two methods of calculating fluxes.*

At low redshift, it can be seen that there is a residual error in the calculated intrinsic far-UV flux even for $N_{\rm max} = 4$. This arises from stars of age about 1 Gyr (i.e. the TP-AGB population), for which the tabulated UV fluxes in the Maraston (2005) population synthesis tables show a large jump in luminosity between the two lowest-metallicity bins. Merging galaxies that contain stars of differing metallicity can therefore lead to large changes in flux. Such merging can occur for any choice of $N_{\rm max}$ and is hence a fundamental (albeit very minor) limitation of the SFH magnitude-reconstruction method.†

The addition of dust significantly degrades the agreement between the two reconstruction methods in all bands, with the far- and near-UV being most affected. At $z = 2$ the rms differences between dust-corrected magnitudes are approximately 0.20 for far- and near-UV, 0.10 for u and g, 0.04 for z and J, and 0.02 for K_s and $\mathrm{irac}_{3.6\mu\mathrm{m}}$.

In the current version of L-Galaxies a two-component dust model applies extinction separately from the diffuse interstellar medium and from molecular birth clouds (see Section 1.14 in the supplementary material of Henriques et al. 2014 for details). The large differences seen for dust corrected magnitudes are mostly caused by the latter. The calculation of this optical depth includes a random gaussian term that leads to differences in the amount extinction assumed for each individual galaxy when a different number of timesteps are used.

The method successfully tested in this section is adopted in the recent major release of the Munich model, HWT14. By computing emission properties in post-processing, the memory consumption of the code is no longer dependent on the number of photometric bands. Moreover, the method allows emission properties to be computed after the model is completed using any stellar populations synthesis code and for the filters used by any

*We do not use a finer time resolution as a default as this agreement is illusory – the underlying merger tree is not capturing the dynamics on that short a timescale.

†The limitation could be overcome by keeping SFHs for several different metallicity bins, but this moves away from the spirit of the method and we do not think that the gain justifies the extra storage cost.

observational instrument. To show the potential of the new method, the new major release already includes emission in 20 bands in the snapshot catalogues, and in 40 bands and for two different stellar populations for the lightcones, all calculated in post-processing.

2.5 Comparison of VESPA and L-Galaxies

For reasons described in Appendix 2.8, we use the version of the VESPA catalogue that was created using the population synthesis model of Maraston (2005) with a one-component dust model.

Note that the VESPA data and the SA models produce output with very different binning. The time resolution in the observations is necessarily very coarse at high redshift, whereas there is no such restriction in the models. In Section 2.5.1 below, we re-bin the model predictions to match those of VESPA; throughout the rest of the paper, we will keep the actual binning returned by the models so as to allow a clearer understanding of the growth of galaxies at high redshift.

2.5.1 The main galaxy sample

The main SDSS galaxy sample covers a redshift range of $0 < z \lesssim 0.35$. In Fig. 2.3 we show the mean SFH for all galaxies using the maximum VESPA resolution of 16 bins. For most galaxies, the data quality is not good enough to independently measure masses in all 16 bins and so the VESPA algorithm will return solutions on bins of varying and lower-resolution width, as described in TWH09. The assumed star formation rate (hereafter, SFR) within each bin depends on its width (it is constant in high-resolution bins, and exponentially decaying in low-resolution bins). Choosing a SFR within a bin is an unavoidable part of the process of parametrizing a galaxy. We have however checked that our conclusions remain unchanged if we: (i) use only galaxies with high-resolution bins, and (ii) use a constant SFR in wide bins.

For each VESPA galaxy, we select the model galaxy that most closely matches it in mass and redshift, then use this to construct a mean SFH. The result is shown in Fig. 2.3 along with the predictions from two versions of L-Galaxies: GWB11–Guo et al. (2011, magenta diamonds and dot–dashed line) and HWT14–Henriques et al. (2014, green squares and dashed lines).

In this and subsequent plots, Δf is defined as the fraction of stars (i.e. the specific stellar mass) within each bin. Unless mentioned otherwise, Δf is calculated separately for each galaxy and then averaged (i.e. weighted by galaxy number rather than stellar mass).

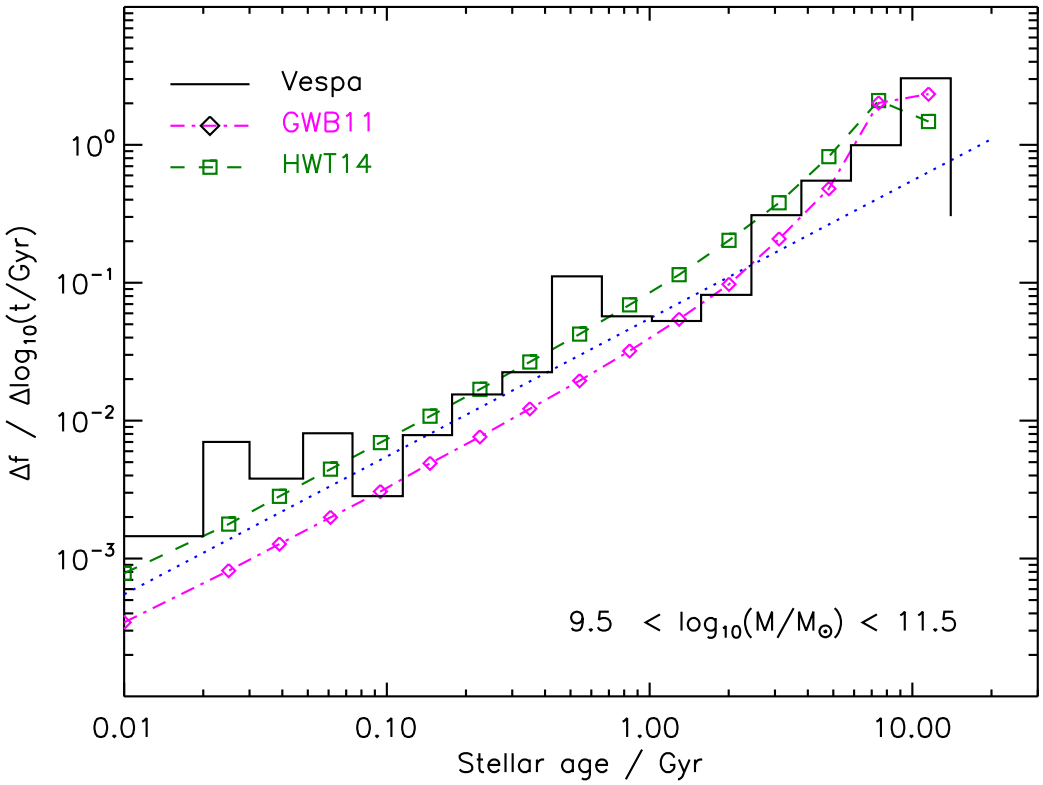

Figure 2.3: The average SFH from VESPA (black lines), the GWB11 model (magenta, dash-dotted lines) and the HWT15 model (dashed, green lines) within the indicated mass range. To guide the eye, the blue, dotted line is the same in each figure and has slope unity, corresponding to a constant SFR.

The VESPA results are not as smooth as the SA model data. We expect the averaged rest-frame SFH of a varied ensemble of galaxies to be devoid of significant structure, and to appear smooth as seen on the model galaxies. At late times, especially, the SFR should be approximately constant, and so one would expect the SFH bins to run parallel to the dotted line, of slope equal to unity. The features seen on the data are heavily dependent on the modelling (see e.g. TWH09, Tojeiro et al. 2013, Appendix 2.8), and therefore are the likely result of limitations of the stellar population synthesis and dust modelling.

The SA models show a turn-down in SFR in the oldest stellar age bin, corresponding to the onset of star formation. No such feature is seen in the VESPA results, most probably because the spectral signatures are too weak to be detected and so the reconstruction method imposes a declining SFR even at these earliest times.

Other than that, at ages above 1 Gyr the GWB11 model seems to provide a reasonable match the observations, whereas for younger stars, the HWT14 model is a better fit. To draw more definitive conclusions about which is the preferred model, one would have to look in much more detail at the reconstruction biases that may be present in the VESPA method when applied to imperfect data, and that will be the subject of future work.

From here on we will rebin the VESPA solutions to five bins in age, to reduce the scatter and to average over these features that we know to be unphysical.

2.5.2 Mass selection

Fig. 2.4 shows the SFHs broken down by stellar mass. The VESPA reconstruction gives a slope that is too steep for high-mass galaxies and too shallow for low-mass galaxies, when compared to the expected constant SFR at recent times. Nevertheless, it can be seen that both the VESPA galaxies, and those in the SA model, form stars earlier in higher-mass galaxies and have a correspondingly lower SFR at late times.

The variation of the model SFHs with mass is explored further in Section 2.6.3, below.

2.5.3 Colour selection

Next, we look at the distinction between red and blue galaxies by selecting according to $u - r$ colour. As HW14 showed, in spite of reproducing the observed galaxy colour bimodality, L-Galaxies does not reproduce the exact colour distributions seen in SDSS. Therefore, applying the same colour cuts in the data and simulation would result in picking out intrinsically different galaxy populations. Instead, we select the 10% bluest and reddest galaxies, according to $u - r$ colour, in both the VESPA and L-Galaxies samples. The

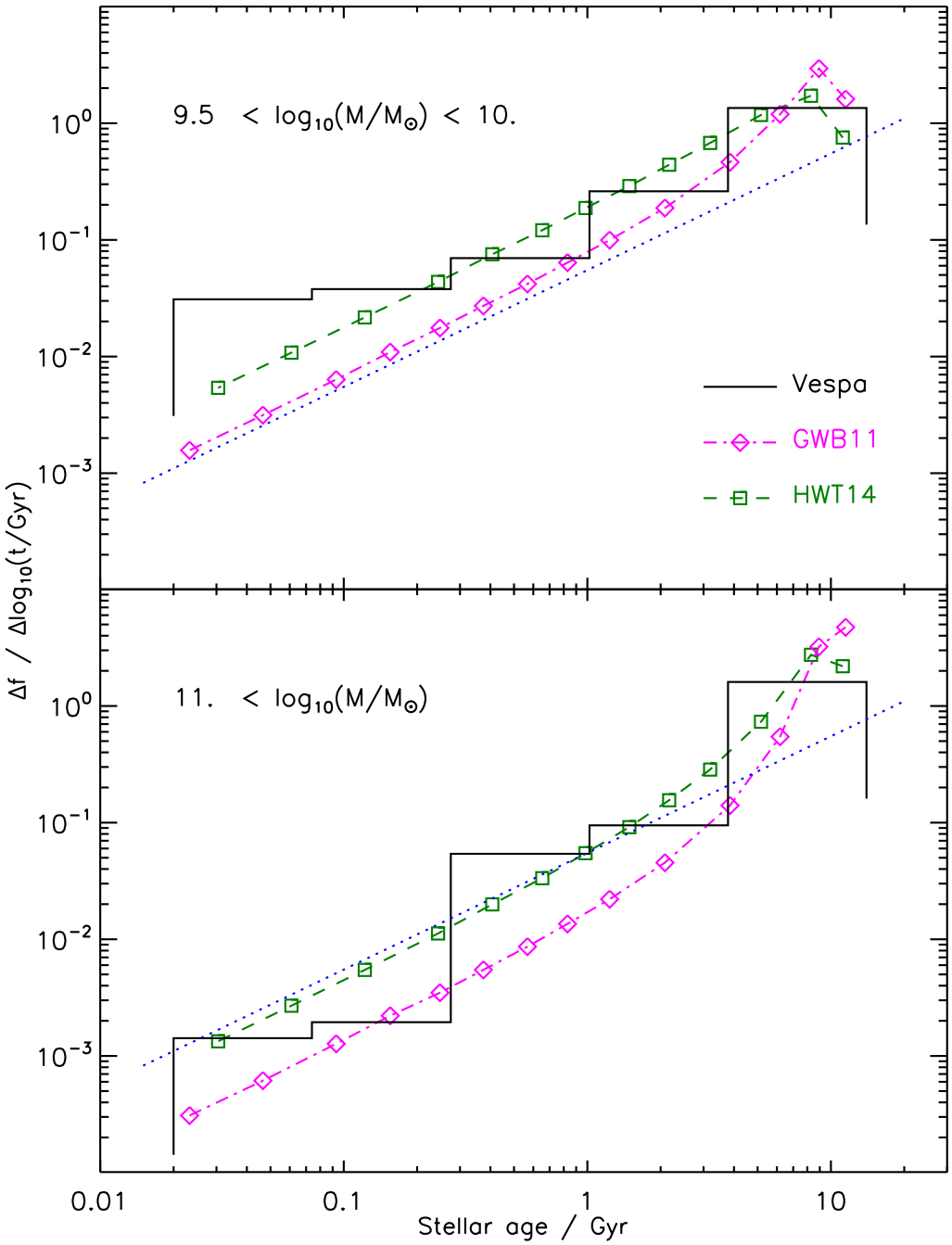

Figure 2.4: The average SFH from VESPA (black lines), the GWB11 model (magenta, dash-dotted lines) and the HWT14 model (dashed, green lines) within two different stellar mass bins, as shown. To guide the eye, the blue, dotted line is the same in each panel and has slope unity, corresponding to a constant SFR.

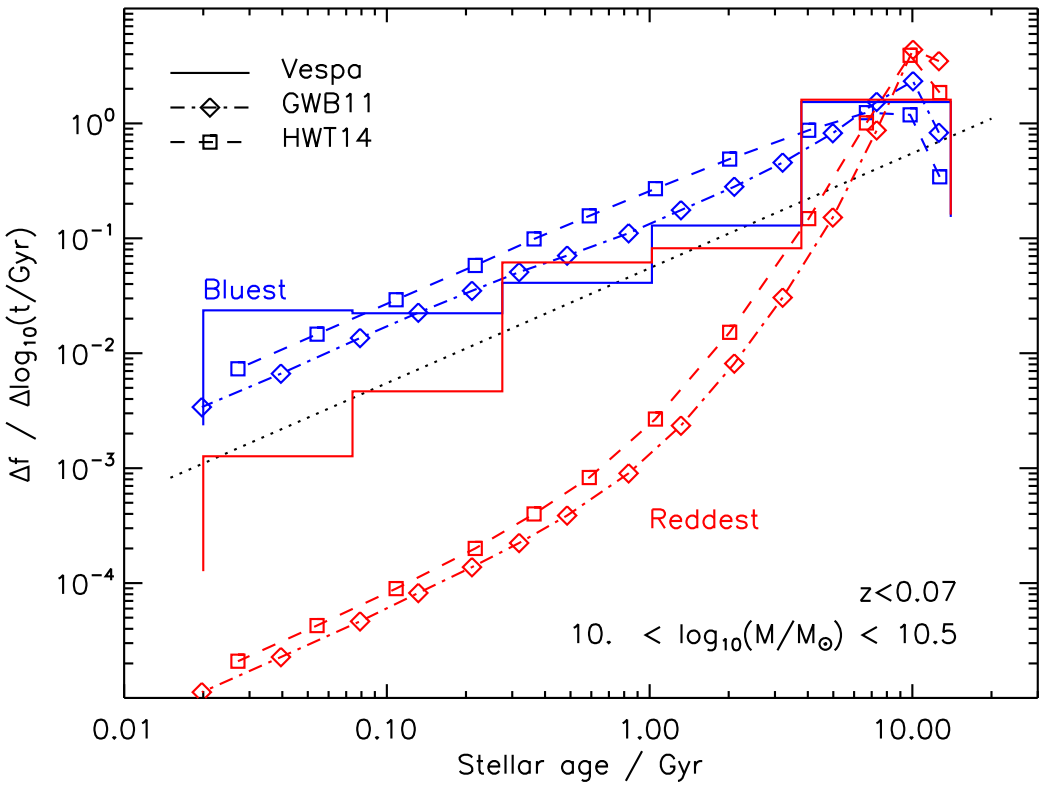

Figure 2.5: The averaged SFH from VESPA (solid lines), the GWB11 models (dashed-dotted lines), and the HWT14 model (dashed lines) for red and blue galaxies with masses greater than 10^{10} and smaller than $10^{10.5}$ $M_{\odot}$ and within redshift interval $0 < z < 0.07$.

resulting SFHs are shown in Fig. 2.5 for galaxies of mass $10^{10} < M/\,\mathrm{M}_{\odot} < 10^{10.5}$ at low redshift, $z < 0.07$.

VESPA produces similar SFHs for both blue and red galaxies except for the youngest stars of age less than 3×10^{8} yr. At first sight, it seems surprising that the deviation between the two populations can have occurred so recently. One interpretation is that galaxies in this mass range may transition back and forth between star forming and quiescent (i.e. show bursts of star formation) on time-scales of this order, and that might also help to explain why the SFR of the bluest galaxies seems to increase to the present day. However, this seems at odds with the observation that red and blue galaxies are observed to have very different metallicities: the stellar-mass-weighted metallicity of the young (age less than 2.5 Gyr) stars is 0.036 in red galaxies,* and 0.019 in blue galaxies, which would suggest that the two form distinct populations.

Whatever the interpretation of the observations, it seems unlikely that they can be made compatible with the model galaxies, which show widely divergent SFHs for red and blue galaxies for stars younger than 5 Gyr. It would seem that termination of star formation is too abrupt in the models as compared to the observations, and lacks the possibility of retriggering of star formation at later times.

The comparison of observed and model galaxies is complicated by the effects of metallicity, dust attenuation and finite fibre aperture on the measured colours. In the models, there is a very strong correlation between stellar mass and metallicity, and between SFR and extinction; in addition; there is no aperture correction. In real galaxies, the scatter is observed to be much higher, and there will be a redshift-dependent colour correction for the finite aperture. This, together with observational error, is likely to move the observed blue and red populations towards each other, so caution should be exercised before drawing definitive conclusions. On the other hand, SA models struggle to match even the colour distribution of galaxies (see, for example, Fig. 9 of HWT14), and the distinction between the models and the observed SDSS data is so large that it is hard to dismiss it lightly. This issue will be investigated in a subsequent paper, and highlights the power in comparing fully-resolved SFHs between models and data.

*In units of the mass fraction of metals with respect to Hydrogen; in these units solar metallicity is $Z_{\odot} = 0.02$

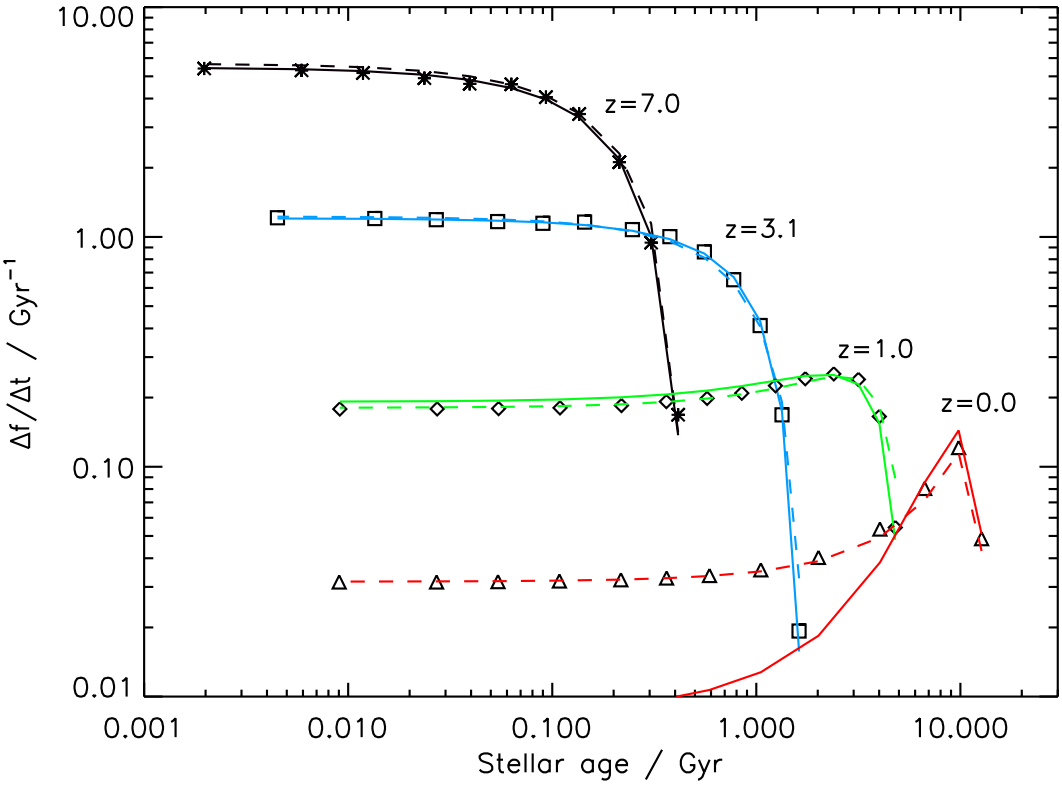

Figure 2.6: The average star formation rates of model galaxies from the HWT15 model, with mass greater than $10^{9.5}\,M_\odot$, at four different redshifts, as shown. The symbols show the model predictions; the curves are fits to the data as described in the text: solid lines, gamma model (Equation 2.1); dashed lines, two-power model (Equation 2.2).

2.6 The evolution of SFHs

This section looks at the predicted SFHs of galaxies at different redshifts. We are interested in both the history of the mean (and median) galaxy population and of the scatter about that mean. This can have important implications for the interpretation of high-redshift galaxies that often rely upon postulated SFHs (Boquien, Buat & Perret, 2014; Pacifici et al., 2015).

2.6.1 Mean SFHs

Fig. 2.6 shows the mean, mass-weighted SFRs of all galaxies with mass greater than $10^{9.5}\,M_\odot$ at four different redshifts: 0, 1, 3 and 7. (Note that this plot differs from previous ones in that we are plotting $\Delta f/\Delta t$ rather than $\Delta f/\Delta \log t$; we do this to make it easier to detect any decrease in the SFR at recent times). The SFR increases rapidly

at early times, then slows down, with a decline to late times (i.e. low stellar ages) being apparent for $z \lesssim 2$.

To characterize the SFR at high redshift, we fit a gamma model featuring a power-law increase in star formation at early times followed by an exponential decline:*

$$\frac{\mathrm{d}f}{\mathrm{d}t} = Ax^p e^{-x}, \quad x = \frac{t_\mathrm{a} - t}{\tau}. \tag{2.1}$$

Here t_a is the age of the galaxy, p sets the rate at which star formation builds up, and τ is the characteristic time-scale over which star formation declines.

At all redshifts above $z = 1$, the SFHs are well-fit by a single set of parameters, $p = 1.5$ and $\tau = 2.0\,\mathrm{Gyr}$, with only t_a varying to reflect the age of the galaxy.† That this is the case is not surprising but reflects the fact that the majority of stars in the Universe are born within galaxies whose mass exceeds 3×10^9 $\mathrm{M}_\odot$: each of the SFHs shown in Fig. 2.6 then mirrors the cosmic SFH. Using this set of parameters then star formation begins in our model at $z \approx 12$, 0.4 Gyr after the big bang, and levels off (i.e. $\mathrm{d}^2 f/\mathrm{d}t^2 = 0$) 3 Gyr later, at $z \approx 2$.

At lower redshifts, it becomes apparent that an exponential decline is too steep. Instead, a two-power model is preferred:

$$\frac{\mathrm{d}f}{\mathrm{d}t} = \frac{A}{x^{-p} + x^q}, \quad x = \frac{t_\mathrm{a} - t}{\tau}. \tag{2.2}$$

Taking $p = 1.39$, $q = 1.33$ and $\tau = 3.0\,\mathrm{Gyr}$ gives a good fit at all redshifts.

Behroozi, Wechsler & Conroy (2013) found identical fitting formulae to those used here to be good fits to the SFHs of galaxies in their abundance matching method to populate haloes with galaxies that match observed stellar mass functions and SFRs. However, for 10^{12} $\mathrm{M}_\odot$ haloes, they find a value of q, that determines the rate of decay of the SFR at late times, to be significantly higher than that quoted above: the reason for this difference is not clear.

We stress that the curves shown in Fig. 2.6 are for the mean star formation rate averaged over all galaxies with mass greater than $10^{9.5}\,\mathrm{M}_\odot$. As is apparent from Fig. 2.4, high mass galaxies form their stars earlier, and low mass galaxies later, than the mean trend. We show in Section 2.6.3 that this is driven primarily by a mass-dependence in the cessation of star formation, and in Section 2.6.4 that there is considerable variation between individual galaxies.

*The same functional form with $p = 1$ was shown by Simha et al. (2014) to be a good fit to the individual SFHs of most of their galaxies in SPH simulations of galaxy formation.

†For $z \gtrsim 2$ these parameters are degenerate, but we choose to freeze them at the values found at lower redshift.

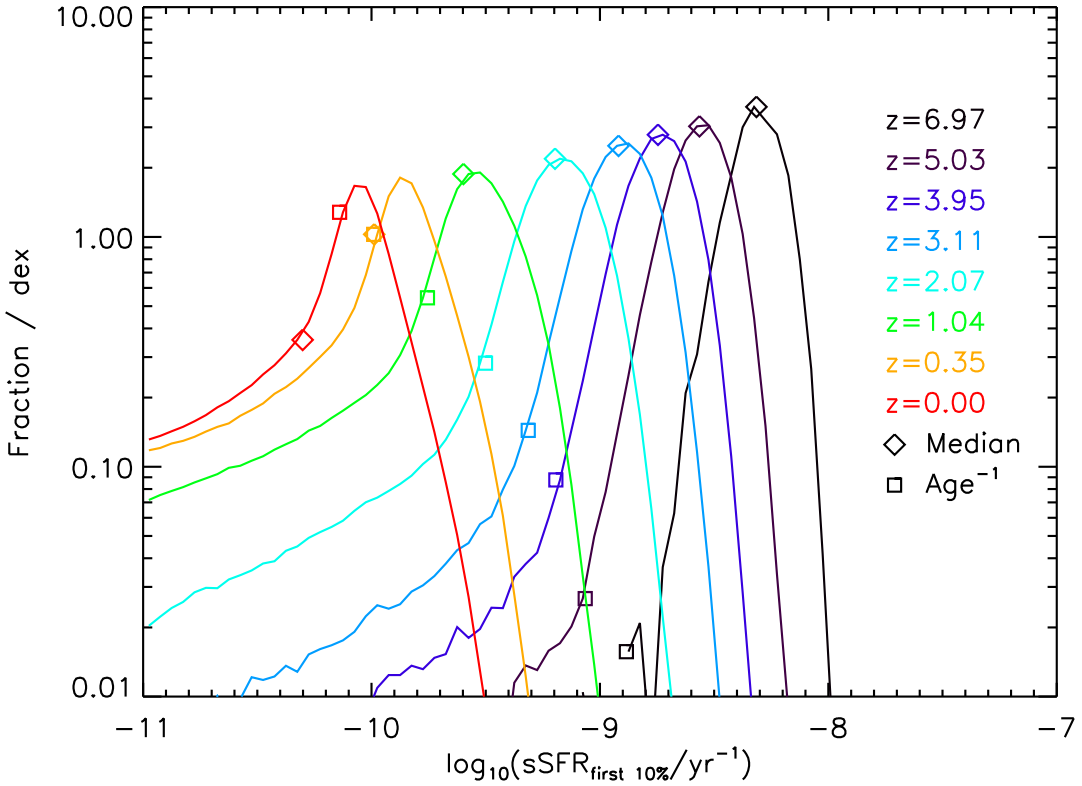

Figure 2.7: The sSFR at different redshifts for the HWT15 model. The subscript "first 10 %" refers to an average over the most recent 10 per cent of the age of the Universe at that time. The diamonds show the median values and the squares show the inverse age of the Universe at each redshift.

2.6.2 Specific SFRs and quiescent fractions

Fig. 2.7 shows the specific SFR (sSFR) of all galaxies with $M > 3\times10^9$ $M_\odot$ in the HWT15 mode at several different redshifts. The SFR here is averaged over the most recent 10 per cent of the age of the Universe at that time. The diamond symbols show the location of the median values, and the squares show the inverse age of the Universe at that redshift.

There is a small spread around the modal sSFR, with the vast majority of star forming galaxies lying within about ±0.25 dex of the peak. However, there is a long tail of galaxies extending to low sSFRs, which becomes more prominent at low redshifts; indeed, at $z = 0$, 35 per cent lie off the left-hand edge of the plot altogether. That is why, below $z = 1$, the median values lie well to the left of the mode.

Above $z \approx 0.35$, most galaxies are forming stars at a rate that would more than double their mass in the age of the Universe;* below that redshift, the opposite is true. It is at

*Note that this is a number-weighted average, so that does not mean that mean SFR peaked at that

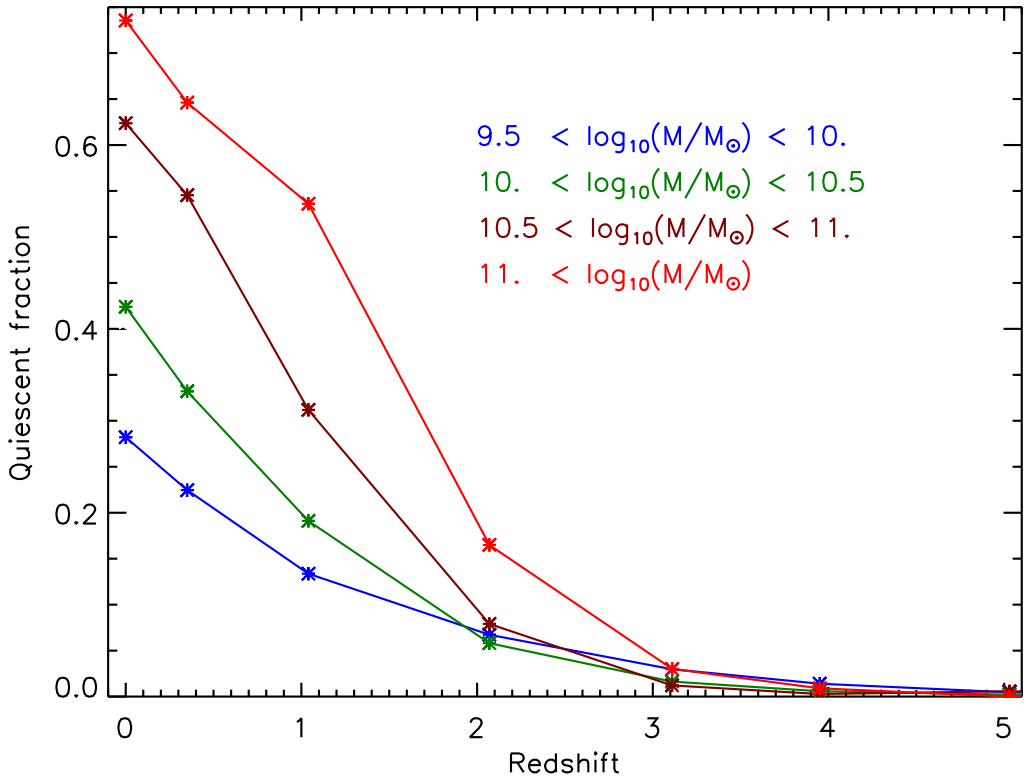

Figure 2.8: The quiescent fraction of galaxies in the HWT15 model, in several different mass bins, and at several different redshifts, as shown.

this time that there is a strong shift from star forming to non-star forming galaxies. There is no sharp distinction between the two but, using Fig. 2.7 as a guide, we define a galaxy to be *quiescent* if it has formed fewer than 3 per cent of its stars in the most recent 10 per cent of the age of the Universe, t_z, at that redshift, i.e. $\mathrm{sSFR}_{\mathrm{first\,10\%}} < 0.3/t_z$.

Fig. 2.8 shows the fraction of quiescent galaxies as a function of mass and redshift. At high redshift, there is some suppression of star formation in dwarf galaxies, reflecting the strong feedback from supernove in the model. However, only a small fraction of galaxies are affected and there is a much greater growth of the passive population at redshifts below $z \approx 3$. Once again, there is clear evidence of down-sizing in that more massive galaxies start to become quiescent earlier than lower-mass ones.

Table 2.3 lists several measures of star formation activity. $S_{0.5,\mathrm{all}}$ is the median sSFR of the sample multiplied by the age of the Universe at that redshift; likewise $S_{0.5,\mathrm{sf}}$ is the same thing, but restricted to star forming galaxies. The two begin to differ significantly below a redshift of about 3 once the quiescent fraction begins to rise. This fraction is

time.

Table 2.3: The median sSFRs, and the quiescent fraction of galaxies in the HWT15 model with mass exceeding 3×10^9 $M_\odot$. The columns are: redshift; age of the Universe in Gyr, t/Gyr; median specific SFR multiplied by the age of the Universe, $S_{0.5,\mathrm{all}}$; the same but restricted to star forming galaxies, $S_{0.5,\mathrm{sf}}$; quiescent fraction weighted by number, Q_{num}; and quiescent fraction weighted by mass, Q_{mass}.

Redshift	t/Gyr	$S_{0.5,\mathrm{all}}$	$S_{0.5,\mathrm{sf}}$	Q_{num}	Q_{mass}
6.97	0.77	3.71	3.71	0.00	0.00
5.03	1.16	3.17	3.17	0.00	0.00
3.95	1.56	2.81	2.82	0.01	0.01
3.11	2.06	2.50	2.53	0.03	0.02
2.07	3.17	2.02	2.09	0.07	0.08
1.04	5.69	1.43	1.60	0.18	0.28
0.35	9.79	0.99	1.31	0.33	0.48
0.00	13.80	0.69	1.17	0.41	0.58

listed weighted both by galaxy number and by galaxy mass, from which it can be seen that about 58 per cent of stars in the current-day Universe lie in galaxies that are not actively star forming.

The medians listed in the table show that there has been a steady decline in star formation activity in galaxies from a redshift of at least 7 right through to the current day. However, even as recently as $z \approx 0.35$, most galaxies were still forming stars at a rate that would more than double their mass within the age of the Universe at that time.

2.6.3 The cause of mass-dependent SFHs

The top panel of Fig. 2.9 shows the SFHs of galaxies at $z = 0$ split into four different mass bins. In order to better illustrate the onset of star formation, which is poorly resolved at $z = 0$ using our default number of bins, for this section only we use $N_{\mathrm{max}} = 4$ (see Section 2.3.3), giving 29 time-bins at $z = 0$. Age downsizing is clearly visible with more massive galaxies forming their stars earlier than lower-mass ones.

To better understand the cause of this, the middle and lower panels show the same curves for star forming, and for quiescent galaxies, respectively. Although the correspondence is not perfect, the agreement between the different mass-bins is much tighter in the central panel than in the upper one. That strongly suggests that the SFHs of star

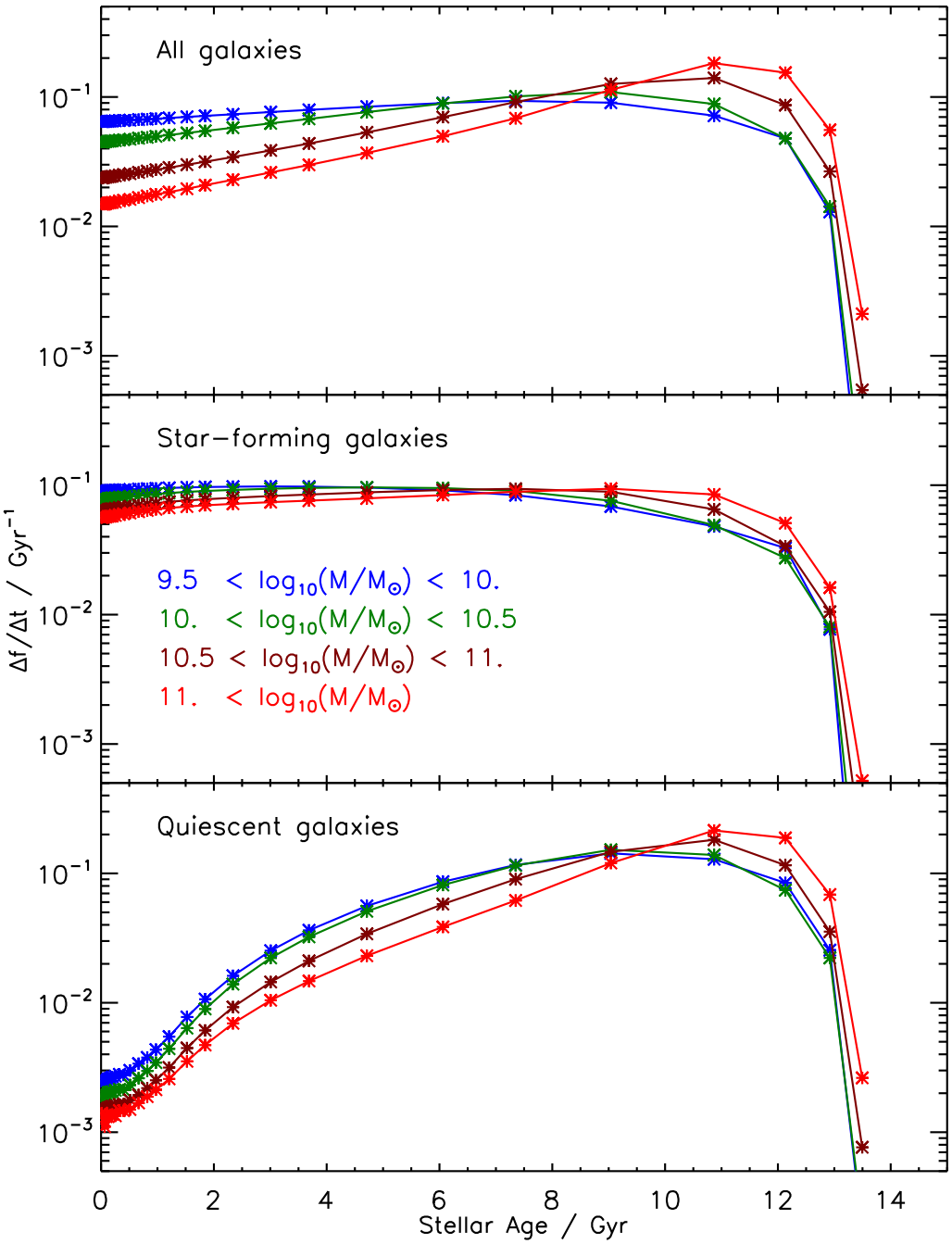

Figure 2.9: The SFR as a function of time for the HWT15 model at $z = 0$, split into 4 different mass bins, as shown. The upper panel shows all galaxies in each mass-range; the middle panel shows star forming galaxies; and the lower panel panel shows quiescent galaxies, as defined in Section 2.6.2.

forming galaxies are very similar, independent of the mass of the galaxy, and that the primary driver of the mass-dependence is the different evolution of the quiescent fraction. Note that we tried only a single definition of quiescence and it is likely that the residual mass-dependence in the central panel could be reduced even further if we optimised the definition for that purpose.

2.6.4 Individual SFHs

In this section, we restrict our attention to star forming galaxies.

Although the mean SFH is well-described by a simple functional form, Fig. 2.10 shows that individual galaxies have a wide variety of histories. This figure shows histograms of the rate of decline of the SFR measured by the ratio in two successive time-bins: the most recent 10 per cent, and the next most recent 10 per cent, of the age of the Universe at that redshift. Galaxies to the left/right of the vertical line have declining/increasing SFRs, respectively.

The distribution of ratios is shown in Fig. 2.10 for a variety of redshifts. There is a gradual shift from increasing to decreasing specific star formation rates as the Universe ages. When measured in this way, an equal balance between increasing and decreasing SFRs is achieved somewhere between redshifts 1 and 2. At all times, however, there is a significant fraction of galaxies lying in each of these populations: at $z = 7$, four-fifths of galaxies show an increasing SFR, and at $z = 0$ three-quarters show a decreasing one.

2.7 Conclusions

In this paper, we have introduced the recording of star formation histories (SFHs) in the L-Galaxies SA model. At any given point in a galaxy's evolution, the mass of recently-formed stars is recorded in bins of time resolution equal to that of the timestep in the SA model (1-2$\times10^7$ yr). These bins are gradually merged together as the galaxy ages, such that older stars are grouped together into larger bins.

We investigate the extent to which SFHs may be used to reconstruct stellar spectra in post-processing; we compare our SFHs to those in the publicly-available VESPA catalogue extracted from SDSS-DR7 data; and we investigate in our favoured SA model (Henriques et al., 2014) the evolution of SFHs as a function of galaxy mass. Our key results are as follows:

- Post-processing reconstruction of magnitudes in various observational bands gives good agreement with on-the-fly accumulation of luminosity, provided that $N_{\rm max} \geq 2$

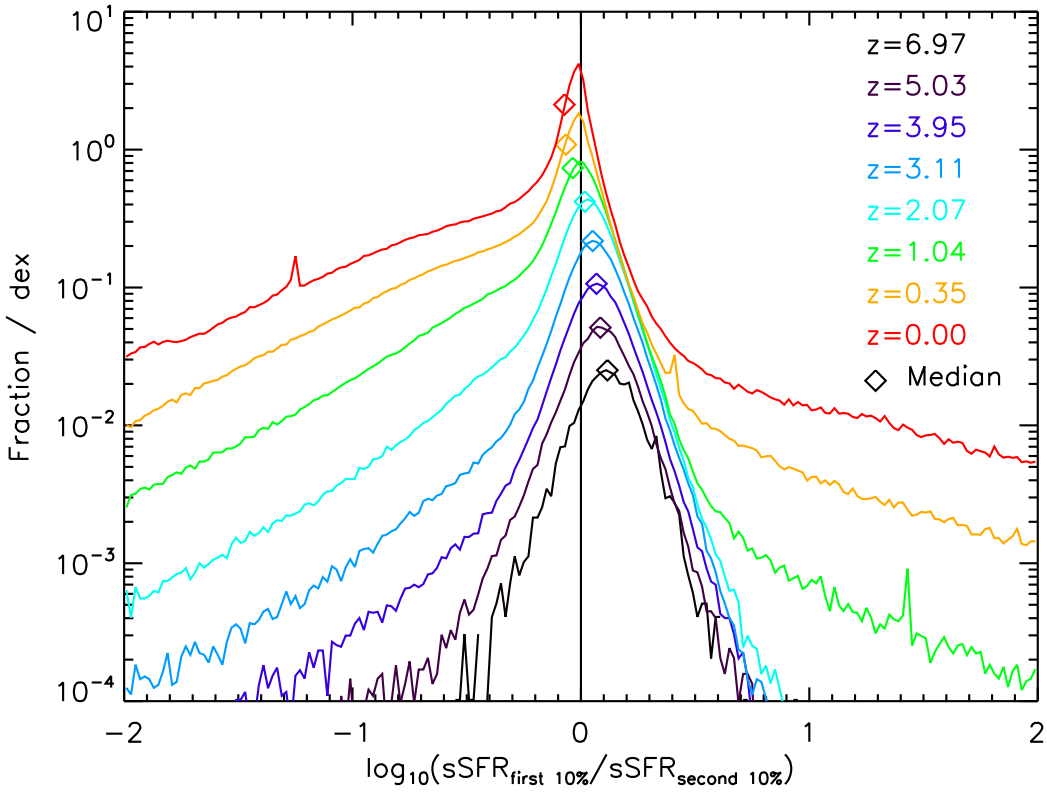

Figure 2.10: The ratio of the SFR in the most recent ('first') 10 per cent, to that in the previous ('second') 10 per cent, of the age of the Universe at that redshift, for galaxies with masses exceeding $10^{9.5}\,\mathrm{M}_{\odot}$in the HWT15 model. The y-scale corresponds to the $z = 0$ curve: the higher-redshift curves are offset by successive factors of two to space them out in the y-direction. The diamond symbols show the median values.

(which equates to 16 bins at $z = 0$). Quantitatively, the rms difference between raw and reconstructed magnitudes is less than 0.05 for all bands except the far-UV, for which it is 0.29.

- The SA models show reasonable qualitative agreement with the observed SFHs of the SDSS Main Galaxy Sample from the VESPA catalogue, with the GWB11 model fitting better for stars older than 1 Gyr, and the HWT15 model fitting better for younger stars.

- When divided by mass, both the observations and models show a trend for more massive galaxies to form their stars earlier and have lower current sSFRs than lower-mass galaxies.

- When divided by colour, the agreement is poorer. Both versions of the SA model show much more extreme variation in SFH with colour than do observed galaxies from the VESPA data base. In the model, the SFHs of red and blue galaxies begin to differ as long ago as 5 Gyr, compared to just 0.3 Gyr for observed galaxies. One possible explanation could be that real galaxies show repeated episodes of star formation that are not present in the models. We note, however, that a more rigorous investigation of the data is required before drawing any definitive conclusions.

- At $z \geq 1$ the mean SFR of all model galaxies with stellar mass greater than $3{\times}10^9$ $\mathrm{M}_\odot$ is well fitted by the formula $\mathrm{d}f/\mathrm{d}t \propto x^{1.5}e^{-x}$, where $x = (t_\mathrm{a}{-}t)/2.0\,\mathrm{Gyr}$. Here t is the lookback time and t_a is the age of the galaxy. At later times, the SFR declines less rapidly and a two-power model (that contains an extra parameter) is a better fit over the whole of cosmic history: $\mathrm{d}f/\mathrm{d}t \propto 1/(x^{-1.39} + x^{1.33})$, where $x = (t_\mathrm{a} - t)/3.0\,\mathrm{Gyr}$.

- Although star formation rates have been declining for more than half the history of the Universe, the typical (median) star forming galaxy today is still forming stars at a rate that will more than double its mass in a Hubble time.

- We define a galaxy to be quiescent if it forms fewer than 3 per cent of its stars in the most recent 10 per cent of the age of the Universe, t_z at that redshift, i.e. sSFR$< 0.3/t_z$. Then the quiescent fraction begins to increase rapidly below $z \sim 3$, reaching 41 % by number and 58 % by mass for galaxies with $M > 10^9$ $\mathrm{M}_\odot$ at the current day.

- Our model produces a small fraction of quiescent dwarf galaxies, $M < 10^{10}$ $\mathrm{M}_\odot$ at all times, but the main effect that we see is consistent with downsizing in that more

massive galaxies become quiescent first, followed by successively lower mass galaxies as the Universe ages.

- When split by mass, and using a finer time resolution in the SFHs, downsizing is very clear to see, with the peak of the SFR shifting from a lookback time of about 11 Gyr in the most massive galaxies (current-day mass greater than 10^{11} $M_\odot$) to less than 8 Gyr in lower-mass systems (3×10^9–10^{10} $M_\odot$).

- When split into quiescent and star forming populations, the differences between the mean SFHs of star forming galaxies of different mass is much reduced. Downsizing thus has its origin in an earlier transition from star forming to quiescent status in galaxies that are more massive at the current-day.

- Although the mean SFHs are well defined, there is a huge dispersion in the SFHs of individual galaxies such that, even at the current day, many galaxies still have increasing SFRs. At $z = 7$, four-fifths of galaxies show an increasing SFR, and at $z = 0$ three-quarters show a decreasing one; an equal balance between galaxies with increasing and decreasing SFRs is achieved somewhere between redshifts 1 and 2.

As can be seen from the above, one of the key drivers of galaxy evolution is the rate at which star formation is quenched. This is investigated in a companion paper, Henriques et al. (in preparation), that undertakes a detailed comparison with observations of the quenched fraction as a function of environment and mass. There it is shown that the HWT15 model does a much better job than previous incarnations of the L-Galaxies SA model in terminating star formation in massive galaxies, whilst allowing continued star formation in low-mass satellites, though the quantitative agreement is still far from ideal.

An earlier paper, Yates et al. (2013), combined the SFHs with a multicomponent model for stellar feedback to investigate the metallicity evolution of galaxies. This then enables us to construct metallicity histories for galaxies along the lines of the SFHs presented in this paper. Unfortunately, the observational data from VESPA is currently unable to constrain the metallicity histories with any degree of certainty.

The low-resolution ($N_{\rm max} = 2$) SFHs for the HWT15T14 SA model, presented in this paper, are publicly available to download from the Millennium data base* and have been used to reconstruct predicted fluxes in post-processing. Higher resolution catalogues are available from the authors upon request.

*http://gavo.mpa-garching.mpg.de/MyMillennium/

Acknowledgements

The authors contributed in the following way to this paper. SS undertook the vast majority of the data analysis and produced a first draft of the paper and figures. PAT supervised SS, led the interpretation of the results, wrote the bulk of the text. BH led and wrote the section on post-processing of magnitudes. RT led the interpretation of the VESPAresults. At different stages each of the other authors (excepting GL) were responsible for discussion of the results, shaping different parts of the paper, and helping to draft the text. GL provided the technical support to help integrate the post-processing of galaxy magnitudes in the Munich GAVO repository (aka the Millennium data base).

This paper draws upon data from the VESPA data base, hosted by the ROE. Much of the data analysis was undertaken on the Cosma-4 supercomputer at Durham and on the Apollo cluster at Sussex.

PAT, SJO and SW acknowledge support from the Science and Technology Facilities Council (grant number ST/L000652/1). RT acknowledges support from the Science and Technology Facilities Council via an Ernest Rutherford Fellowship (grant number ST/K004719/1). GL was supported by Advanced Grant 246797 'GALFORMOD' from the European Research Council and by the National Science Foundation under Grant no. 1261715.

2.8 The choice of VESPA catalogue

The VESPA SDSS-7 catalogue (Tojeiro et al., 2009) contains galaxies with a wide variety of data quality, some showing reconstructed mass errors that are greater than 100 %. In order not to bias the results, we include the whole sample in our analysis. We show average SFHs weighted by galaxy number, rather than galaxy mass, so as to minimize the effect of the errors in the mass reconstruction. We have checked that restricting the analysis to the galaxies with the best data quality does, in fact, lead to qualitatively similar SFHs.

The average SFHs of galaxies in this subsample, weighted by galaxy number, are shown in Fig. 2.11 for two different SEDs (Bruzual & Charlot 2003, hereafter BC03 and Maraston 2005, hereafterM05) and two different dust models. The one-component dust model is a uniform screen applied to the whole stellar population; the two-component model adds in extra absorption in front of young stars.

First note that the M05 models show a much smoother change in the SFR between look-back times of 0.1-10 Gyr than do those of BC03. In such a large galaxy sample, it is

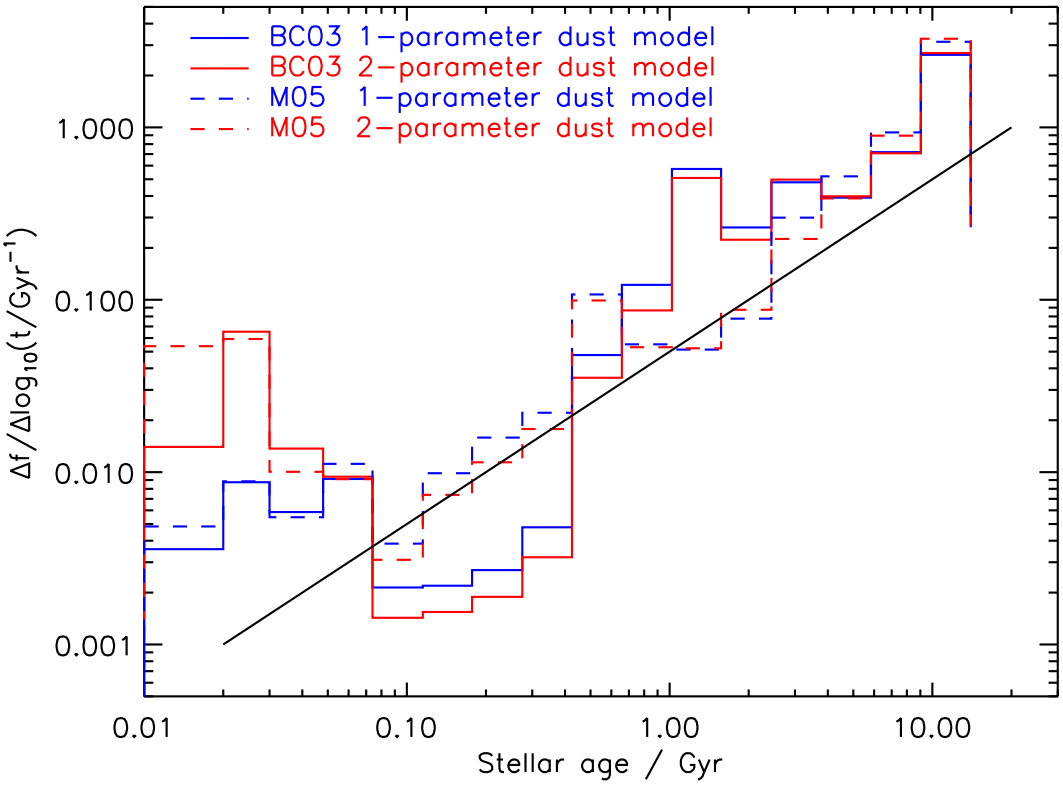

Figure 2.11: The average SFH of VESPA galaxies obtained using the SEDs of BC03 (solid lines) and M05 (dashed lines). Blue lines show results for a one-parameter dust model whereas red shows a two-parameter dust model. The solid, black line has a slope of unity, corresponding to a constant SFR.

hard to think of a plausible reason for this and the SA models show no such feature. For that reason, we use the M05 results.

Both the BC03 and, to a lesser extent, the M05 results for the two-dust model show a significant increase in SFR at ages less than 0.1 Gyr. Again, this seems implausible and suggests that there is not enough constraining power in the data: the model has presumably confused dust obscured, young stars with some older population, perhaps to explain some spectral feature that is not well fitted by the SEDs.

Throughout the body of the paper, we use the VESPA results for M05 and a single dust model.

Bibliography

Angulo R. E., White S. D. M., 2010, MNRAS, 405, 143 Cited on 28

Angulo R. E., White S. D. M., Springel V., Henriques B., 2014, MNRAS, 442, 2131 Cited on 29

Behroozi P. S., Wechsler R. H., Conroy C., 2013, ApJ, 770, A57 Cited on 43

Boquien M., Buat V., Perret V., 2014, A&A, 571, A72 Cited on 42

Boylan-Kolchin M., Springel V., White S. D. M., Jenkins A., Lemson G., 2009, MNRAS, 398, 1150 Cited on 29

Bruzual G., Charlot S., 2003, MNRAS, 344, 1000 Cited on 52

Calzetti D., 1999, Ap&SS, 266, 243 Cited on 26

Cid Fernandes R., Gu Q., Melnick J., Terlevich E., Terlevich R., Kunth D., Rodrigues Lacerda R., Joguet B., 2004, MNRAS, 355, 273 Cited on 27

Cid Fernandes R., Mateus A., Sodré L., Stasińska G., Gomes J. M., 2005, MNRAS, 358, 363 Cited on 27

Cohn J. D., van de Voort F., 2015, MNRAS, 446, 3253 Cited on 26

Guo Q. et al., 2011, MNRAS, 413, 101 Cited on 36

Heavens A., Panter B., Jimenez R., Dunlop J., 2004, Nature, 428, 625 Cited on 27

Henriques B., White S., Thomas P., Angulo R., Guo Q., Lemson G., Springel V., 2013, MNRAS, 431, 3373 Cited on 28

Henriques B., White S., Thomas P., Angulo R., Guo Q., Lemson G., Springel V., Overzier R., 2014, MNRAS, accepted, astro-ph/1410.0365 Cited on 26, 28, 35, 36, 48

Kennicutt, Jr. R. C., 1998, ARA&A, 36, 189 Cited on 26

Koleva M., Prugniel P., Bouchard A., Wu Y., 2009, A&A, 501, 1269 Cited on 27

Lemson G., The Virgo Consortium, 2006, preprint (astro-ph/0608019) Cited on 26

Maraston C., 2005, MNRAS, 362, 799 Cited on 35, 36, 52

Ocvirk P., Pichon C., Lançon A., Thiébaut E., 2006, MNRAS, 365, 46 Cited on 27, 29

Pacifici C. et al., 2015, MNRAS, 447, 786 Cited on 42

Panter B., Heavens A. F., Jimenez R., 2003, MNRAS, 343, 1145 Cited on 29

Panter B., Jimenez R., Heavens A. F., Charlot S., 2007, MNRAS, 378, 1550 Cited on 27

Planck Collaboration XVI, 2014, A&A, 571, A16 Cited on 28

SDSS Collaboration, 2009, ApJS, 182, 543 Cited on 27

Simha V., Weinberg D. H., Conroy C., Dave R., Fardal M., Katz N., Oppenheimer B. D., 2014, astro-ph/1404.0402 Cited on 43

Springel V. et al., 2005, Nat., 435, 629 Cited on 28

Tojeiro R., Heavens A. F., Jimenez R., Panter B., 2007, MNRAS, 381, 1252 Cited on 27, 29

Tojeiro R. et al., 2013, MNRAS, 432, 359 Cited on 38

Tojeiro R., Wilkins S., Heavens A. F., Panter B., Jimenez R., 2009, ApJS, 185, 1 Cited on 27, 52

Yates R. M., Henriques B., Thomas P. A., Kauffmann G., Johansson J., White S. D. M., 2013, MNRAS, 435, 3500 Cited on 27, 51

Chapter 3

Paper2

Matching Herschel galaxy luminosity functions and number counts in the L-GALAXIES semi-analytic model

Sorour Shamshiri, Peter A. Thomas, Seb J. Oliver

3.1 Abstract

s

We extend the L-GALAXIES semi-analytic model to predict the infrared luminosity function and number counts in the SPIRE bands. We use the Kennicutt (1998) relation to convert the model star formation rate (SFR) to infrared luminosity. We adopt the SED library of normal galaxies in which SEDs vary with redshift in order to drive monochromatic flux densities. We compare predictions from the L-GALAXIES semi-analytic model for the far-infrared galaxy luminosity functions and number counts with observations from Herschel. Our model shows lower luminosities at a given number density where this disagreement increases gradually with redshift. Also our model estimates lower number densities at a given flux in comparison to observed data in SPIRE bands.

We consider a model in which the luminosity of individual sources is increased, whilst lowering the number density in proportion. We adopt this model with two different approaches: first we increase the luminosities and decrease the number densities at the given luminosity in order to preserve the total flux. In another approach, luminosity and the number density of individual objects are changed independently, thus the total flux is not preserved. We find that the overall agreement is good between both simulated and observational number counts after fitting with the number count increases with fluxes up

to 0.1 Jy before it declines with fluxes above 0.1 Jy.

After fitting the luminosity functions assuming the conservation of the total flux, our number counts still show some deviations from published values. In contrast, we have a better fit for LF and are able to produce the same number count as observations if number density and luminosity are allowed to scale separately. We show that this can be explained by differences in the luminosity functions used in extraction of the number counts from those that we find in our results.

3.2 Introduction

Star formation (SF) is one of the main physical mechanisms underpinning galaxy formation and evolution. Although observers use a variety of techniques, gaining an accurate estimate of SFR from observational data is always challenging, one major cause of uncertainty being dust extinction. Light from young, bright, blue stars, closely related to the instantaneous SFR of the galaxies, is frequently attenuated by dust and re-radiated at far-infrared (FIR) wavelengths (Puget et al., 1996; Fixsen et al., 1998; Calzetti, 2001), which makes knowledge of the FIR crucial in this context.

The infrared emission has been detected by the ground-based, sub-millimetre Common-User Bolometer Array (SCUBA) on the James Clerk Maxwell Telescope in Hawaii at 450 and 850 microns and resolved into individual sources (Holland et al., 1998; Barger et al., 1998). These sources are commonly extremely luminous, high SFR galaxies at $z > 2$ (Coppin et al., 2008) with significant dust extinction (Chapman et al., 2005). However, the contribution of SCUBA-resolved galaxies is almost 30 times less that the energy density at the 200 micron peak of the FIRB distribution and it is suggested by spectral shape of the FIRB and stacking analyses (Pascale et al., 2009; Dole et al., 2006) that the peak of the FIRB is actually produced by galaxies at lower redshift (Gispert, Lagache & Puget, 2000; Dole et al., 2006; Pascale et al., 2009).

The Herschel Space Observatory was launched in May 2009 and finished its mission in 2013 having vastly improved observations of the Universe in the FIR. With Herschel we are now able to explore the emission of obscured galaxies from 250 to 500 and 70 to 160 microns with the Spectral and Photometric Imaging Receiver (SPIRE, Griffin et al., 2010) and Photodetector Array Camera & Spectrometer (PACS, Poglitsch et al., 2010), respectively. It allowed us to fully constrain the peak of the FIR background and measured the FIR luminosity function of galaxies up to z=4 (Gruppioni et al., 2013; Magnelli et al., 2013), revealing a steep increase from z=0 up to z$\sim$1, then remaining approximately

constant between z∼1 and z∼3 before decreasing at z>3.

In addition, observational measurements of intrinsic physical properties (e.g. stellar mass and SFR) from multi-wavelength photometry are vastly affected by poorly constrained biases and can vary significantly depending on which observational indicator is used to estimate them (Lee et al., 2009). To better constrain the models we need a self-consistent galaxy model that can predict observations across the whole spectral range.

Semi-analytic models (SAMs) of galaxy formation, often built upon the skeleton of a merger tree predicted by cold dark matter (CDM), use simplified physical prescriptions that capture the main processes that govern galaxy formation. SAMs successfully reproduce many observed galaxies properties, such as total stellar mass or luminosity, SFR and chemical enrichment history. Here, we follow the latest Munich SAM, known as L-Galaxies, as described in Henriques et al. (2015), hereafter HWT15, extending its predictions to the FIR.

Over the last decade, SA models have continued to extend their predictions to observed quantities at near/far infrared wavelengths. Particularly, GALFORM versions use the model predictions for the mass and geometrical distribution of the dust and couple SAM with a physical radiative transfer model in order to predict the dust absorption self-consistently (e.g. Baugh et al. (2005); Fontanot et al. (2007); Lacey et al. (2011); Gonzalez-Perez et al. (2013); Lacey et al. (2015)). The early version of GALFORM SAM (Cole et al. (2000)) that employs this sophisticated dust model, underestimates the number of sub-millimetres galaxies and the luminosity function at high redshifts. Baugh et al. (2005) suggested that the only technique to reconcile the model predictions with observations of high-redshift galaxies was to embrace a top-heavy stellar initial mass function (IMF) when stars formed in the starburst triggered by galaxy mergers.

The principle aim of this paper is to determine whether the same results hold in the latest version of L-Galaxies. We extended HWT15 to predict IR luminosity and calculate the number density of the infrared galaxies. The unified model that we propose in this paper uses a simple and efficient way to convert the model-predicted star formation rates (SFRs) to IR luminosity and IR counts. An important feature of our approach is avoiding the uncertainty due to the complexity of calculating the dust absorption using a radiative transfer model.

In Section 3.3 below we describe how we use the L-Galaxies SAM to derive FIR luminosities and number counts, including map-making and source extraction. Section 3.4 presents results for the total flux density. In Section 3.5, we present our prediction of the

bolometric FIR luminosity function and show how our predicted FIR luminosities must be scaled to fit the observational data. Section 3.6 compares our prediction of number counts with the available data sets both before and after map-making and source extraction. Finally, we discuss our results and draw appropriate conclusions in Section 3.7.

3.3 Method

3.3.1 L-Galaxies

Semi-analytic models connect the observed properties of the galaxy population – abundances, scaling relations, clustering and their evolution with redshift – to the astrophysical processes that drive the formation and evolution of individual galaxies. They are much less time-consuming than hydrodynamical simulations and thus allow one to test a much wider range of physical models.

In this paper, we use the latest version of Munich semi analytic model, L-Galaxies, as described in (Henriques et al., 2015, hereafter HWT15). This model is calibrated to reproduce the observed evolution both of the stellar mass function and the colours of galaxies from $z = 3$ to $z = 0$. It also accurately follows the evolution of the mean SFR as a function of mass over that range, except for a slight deficit in SFR for the brightest galaxies at $z = 3$.

L-Galaxies establishes the hierarchical growth of structure from the Millennium Simulation (MR, Springel et al., 2005) adapted to use the *Planck* cosmology (Planck Collaboration XVI, 2014): $\Omega_\Lambda = 0.683$,$\Omega_\mathrm{m} = 0.315$, $\Omega_\mathrm{b} = 0.0488$, $\sigma_8 = 0.826$, $n_\mathrm{s} = 0.958$, $h = 0.673$. This gives a box size of $480.3\,h^{-1}$ Mpc and a particle mass of $9.61 \times 10^8 h^{-1}$ $\mathrm{M}_\odot$. L-Galaxies has also been constructed on the higher resolution Millennium-II simulation (MR2 Boylan-Kolchin et al., 2009) which has the same number of particles as the original MR in a box that has one-fifth of the linear size: hence 5 times greater length and 125 greater mass resolution.

Both MR and MR2 provide data at 58 snapshots separated in time by about $3{\times}10^8$ yr at low redshift (less at earlier times) and that data is used in Sections 3.4 & 3.5. The MR has also been used to construct light-cones and those are used in Section 3.6 to investigate number counts – we show in Section 3.5 that the unresolved sources are too faint to show up in the observations.

3.3.2 Conversion between SFR and FIR luminosity

Our model to obtain total infrared luminosity of galaxies is based on the approach proposed by Kennicutt (1998). As noted before, dust absorbs the bulk of the emitted UV light and re-radiates the energy mostly in the FIR. Since most of the UV emission is produced by young stars, the total FIR luminosity is strongly correlated with the SFR. Thus we directly relate the bolometric FIR luminosity (the integrated infrared luminosity between 8 and 1000 μm) to the total SFR using the equation:

$$L_{\rm FIR}/L_{\odot} = K \times {\rm SFR}/M_{\odot}{\rm yr}^{-1}. \tag{3.1}$$

The conversion factor, K, is a function of, amongst other things, the age of the stellar population. We taken a value $K = 5.8 \times 10^9$ from Kennicutt (1998) multiplied by 1.73 to convert to the Chabrier (2003) IMF (Kriek et al., 2009).

3.3.3 Spectral energy distributions

Flux density, S_ν, is computed by the following formula from Hogg (1999):

$$S_\nu = \frac{(1+z)L_{(1+z)\nu}}{4\pi D_L^2}, \tag{3.2}$$

where D_L is the luminosity distance of the source at redshift z and $L_{(1+z)\nu}$ is the corresponding luminosity at frequency $(1+z)\nu$.

The SED's used in our model originally came from Magdis et al. (2012) SED library, which contains different templates for main-sequence (MS) and starburst (SB) galaxies, getting warmer with redshift. Since our model produces inconsequential numbers of star bursts, we used only the main-sequence (MS) SEDs after a re-interpolation on a finer redshift grid provided in Béthermin et al. (2012a).

The following equation is used to compute observed flux at SPIRE and PACs bands:

$$S_\nu = \frac{(1+z)L_{\rm FIR}T_{(1+z)\nu}}{4\pi D_L^2 \int_{8\mu{\rm m}}^{1000\mu{\rm m}} T_v {\rm d}\nu}, \tag{3.3}$$

where T_ν is the observed flux density of main-sequence (MS) galaxies at the given redshift taken from Béthermin et al. (2012a). We produce catalogues at 70, 160, 250, 350 and 500 μm.

3.3.4 Map-making

Because of the poor angular resolution at long-wavelengths, many resolved sources in the optical and near-infrared become unresolved and blended in the Herschel bands (Nguyen et al.,

2010). As a result, extra care needs to be taken in order to estimate correct number counts. Consequently, different efforts have made to correct the maps for blending, flux boosting and incompleteness (e.g. Crawford et al., 2010; Oliver et al., 2010; Valiante et al., 2016). To make our comparison as accurate as possible, we produced mock Herschel maps and extracted sources from them afterwards. We created the maps using the SMAP software package (Levenson et al., 2010a; Viero et al., 2013) for SPIRE bands (250, 350 and 500 μm) in the COSMOS field. In this technique, the simulated catalogue is turned to the image of sky (map) to feed into the HerMES source detection algorithms. HerMES maps are constructed using the standard HIPE (Herschel Interactive Processing Environment, Ott (2010), Valtchanov et al. (2014)) pipeline.

3.3.5 Source extraction

Fluxes of sources are extracted from the SMAP maps using the SUSSEXtractor algorithm, as described in Savage & Oliver (2007). This method finds the peaks in the image and then models the source and the sky at each possible source position. It results in a catalogue including the measured flux density and background flux density both with associated errors.

3.4 Total flux density

We are able to directly calculate the predicted cosmic infrared background (CIB) for the SPIRE bands by integration over the flux from each of our sources using Equation 3.3. Fig. 3.1 presents the cumulative CIB contribution as a function of flux cut. We should stress that we have also checked the CIB value using MR2 and found a consistent result in respect to MR. Therefore, we present only the result of the original simulation (MR) here.

We find a total contribution of the galaxies to the CIB of 7.2, 6.7 and 5.5 $\mathrm{nW\,m^{-2}\,sr^{-1}}$ at 250, 350 and 500 μm, respectively. This value accords at better than 2σ at 250 and 350 μm and 3σ at 500 μm with the FIRAS absolute measurements implemented by Fixsen et al. (1998) or Lagache et al. (2000), shown by the dashed, purple region in the figure. Also shown on the figure are measurements of the CIB above different flux cuts (Béthermin et al., 2010; Oliver et al., 2010; Béthermin et al., 2012b). We find that although our prediction of total CIB is within 2 σ agreement, we strongly overpredict CIB emitted at flux density cuts above 0.01 Jy.

It should be noted that we have investigated these high flux objects in the following

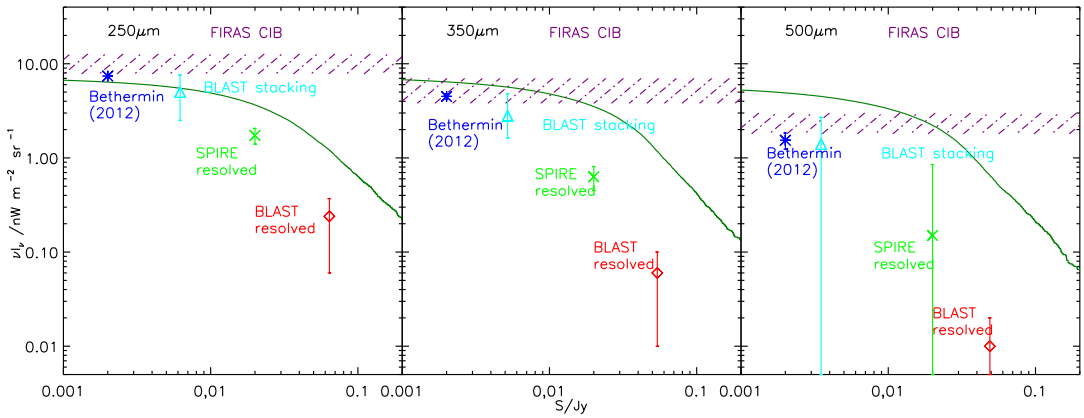

Figure 3.1: Cumulative contribution to the CIB as a function of flux at 250, 350 and 500 μ. The green solid line is cumulative contribution from our predicted number counts. The symbols show: blue asterisk: the fraction resolved at the limit used by Béthermin et al. (2012b); cyan triangle: contribution of the BLAST sources probed by stacking (Béthermin et al., 2010); green cross: the sources resolved by SPIRE (Oliver et al., 2010); red diamond: resolved BLAST sources (Béthermin et al., 2010). The purple, shaded region show the FIRAS absolute measurement of the CIB in the 1-σ confidence region of Fixsen et al. (1998) and Lagache et al. (2000).

sections and discuss the possible explanations behind the inflation of CIB at higher flux density. The CIB contribution can be affected by using different selection of SEDs.

3.5 Infrared Luminosity

3.5.1 Bolometric FIR luminosity function

As described above, SAM model represents its predictions based on two different simulations, MR and MR2, where the two models differ in size and resolution. The original Millennium simulation (MR) has a five-times-bigger box size with much smaller mass resolution in comparison to MR2. In this section, our result is presented using both simulated boxes and are compared with Gruppioni et al. (2013) in several redshift bins, from z $\simeq$ 0 up to z$\sim$4.

Fig. 3.2 compares our predicted FIR luminosity function with that measured by Gruppioni et al. (2013). Note that in both Figures 3.2 and 3.4 the y-axis shows the luminosity density $N \times L_{\rm FIR}$ instead of the number density N: thus the area under the curve traces the total luminosity density. The solid lines are HWT model and diamond

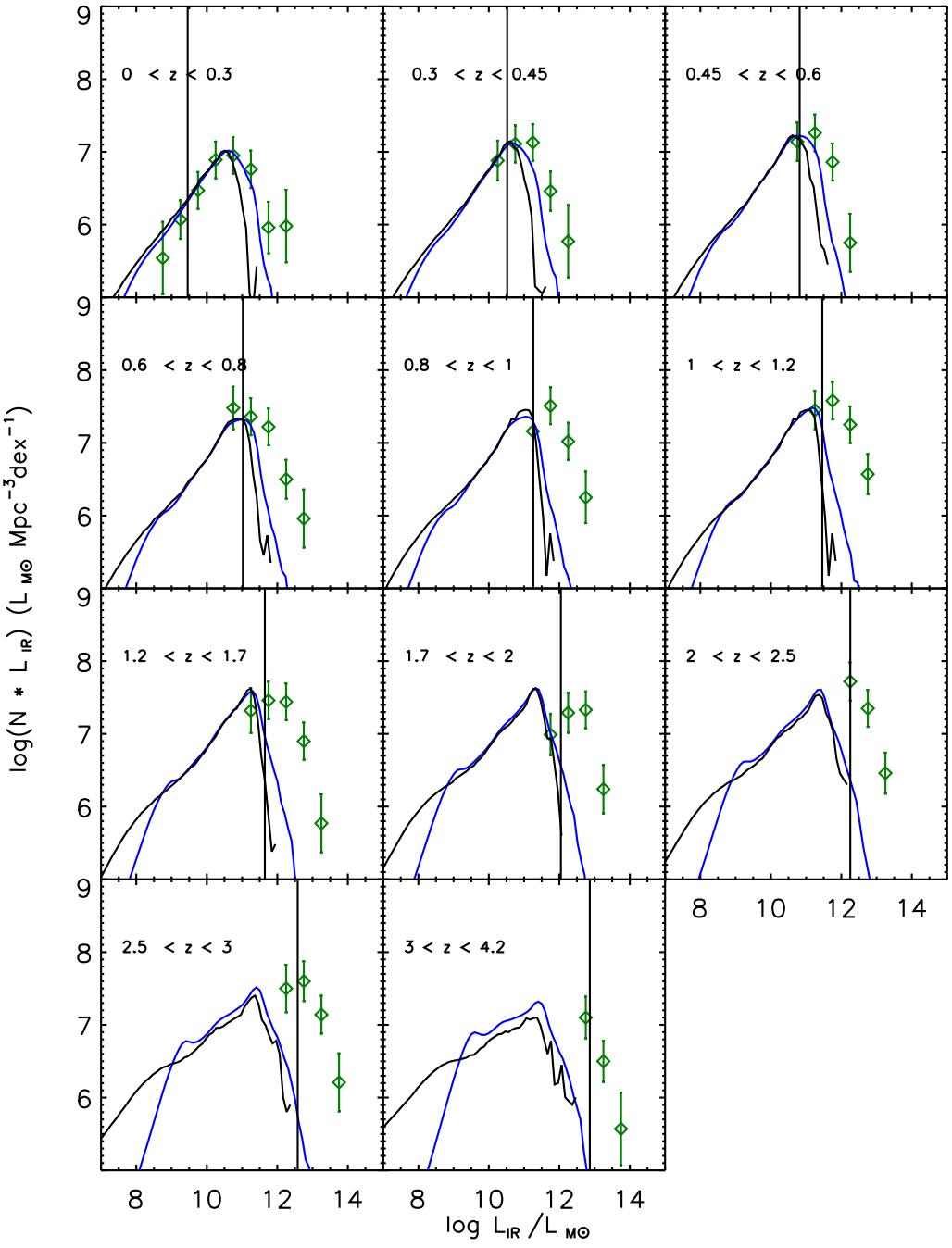

Figure 3.2: Total FIR luminosity density from the L-Galaxies SAM compared to the data from Gruppioni et al. (2013). The blue solid line is using MR and the solid black line is MR2. The vertical line in each panel (except the top-left) shows the luminosity corresponding to a flux limit of 0.01 Jy at the lowest redshift of the bin – for the top-left panel, a redshift of 0.1 is used.

symbols are from Gruppioni et al. (2013).

Infrared luminosities are given by the equation 3.1 which uses the derived SFR from the L-Galaxies model. For this study, we use the result of the snapshots and each of them has a fixed redshift. Therefore, for each redshift bin we use the result of the closest box to the average redshift of each bin.

We would expect that due to the smaller simulated box with higher resolution, MR2 is better in capturing small sources whereas MR fairly represents more massive and rare objects. As a result, we illustrate whether using only MR could plausibly mean that we are missing faint sources. The Fig. 3.2 shows, while MR has a reasonably good alignment with the observation at lower redshifts, MR2 is able to capture sources of lower luminosity than MR. The increased number of less luminous galaxies (at very low fluxes below the data) at MR2 result from the high resolution. Both simulations are unfortunately unable to produce enough luminosity particularly at higher redshifts, which suggests that we need to generate higher luminosities at redshifts greater than 1.5 . The vertical solid black lines reported in each panel represent the luminosity which matches our flux limit at 0.01Jy. Everything on the left of the vertical line has no contribution to the number counts above that flux limit whereas the luminosities on the right are important in the augmentation of number count and flux. In contrast to MR2, the MR simulation retrieves higher luminosities where most of the flux is generated because of the better sampling of massive objects in this cosmological volume. MR2 has a negligible effect for fluxes below 1 mJy and so we use only MR simulation to enable us to retrieve fluxes. Therefore, we now refer SAMs results for using only MR throughout the rest of the paper.

As shown in the Fig. 3.2, our derived total IR LF is a reasonable match to the data points in the first redshift bin, $[0, 0.3]$, however, it moves increasingly away from the observed data as the redshift is increased. The total overall flux is compatible with the observations, but the peak luminosity fails to grow with redshift as rapidly as is observed. The cause of the disparity is unclear, although it may be either because of the lack of high star forming galaxies (or perhaps an under-estimation of the SFR in mergers) or a variation over time of the normalising factor between SFR and IR luminosity or even the combination of both.

Our estimation of IR luminosity is dependent on the conversion factor (Kennicutt, 1998) which relies on IMF, the age of the stellar population and the fraction of light turned to infrared emission. Although we know which IMF we are implementing in our simulation, there is uncertainly in the value of the other two which can result in significant

deviations. Despite the fact that we are underestimating the luminosity at high redshifts, by shifting the peak of the simulated IR LF towards the higher luminosity, we can get a better alignment with observation. We investigate in the next section the degree to which we would have to scale our luminosities to match the observations.

3.5.2 Fitting the luminosity function

Fig. 3.2 shows that at the brighter end of the LF, our model falls bellow the observed one at higher redshifts. We investigate whether we can fit the luminosity functions by a simple approach which allows us to identify where our model needs to be changed and also how this scaling affects the number counts later on.

To characterise the LF, we adopt a two-parameter fitting model which is given by:

$$\phi' = \frac{\beta}{\alpha}\phi(\frac{\mathrm{L}}{\alpha}). \tag{3.4}$$

Here, α and β are the fitting-parameters and ϕ' is the adjusted number density. We consider two different approaches for fitting our model LF to Gruppioni and Bethermin's LF. In order to fit Gruppioni's LF we consider β=1 which means that it preserves the total flux before and after fitting, with the luminosities increasing and number densities decreasing in proportion - in other words our luminosity functions slide to the right on the plot. While in another approach, we treat α and β separately, meaning that not only the model luminosity slides to the right but also the number density at the given luminosity is changed on the plot. Fig. 3.3 contains the value of α and β for each redshift together with a smooth fit to the α values that we use when calculating fluxes for number counts in the following section.

In Fig. 3.4 we compare model predictions before and after fitting to observational estimates of the LF from z = 0 to z = 4. The total IR LF of the original model prediction is plotted with a solid blue line, and the new adjusted-LF by a dashed green line. In general, our curves are relatively aligned up to redshift 0.6 before diverging at higher redshifts. Our scaled curve, when β=1, fits reasonably well to all the data points for all redshift ranges. Also shown in the figure are the model fits from Béthermin et al. (2012b): those match the observations better at z =0.6-0.8 but overestimate the peak of the luminosity function and hence the total flux in some higher redshift bins. We will discuss the effect of this in Section 3.6 below and will show that if the simulated LFs are fitted to Bethermin's LFs without considering the conservation of total flux (by applying the α and β values from Fig. 3.3) then the adjusted number count will pass the Bethermin number count.

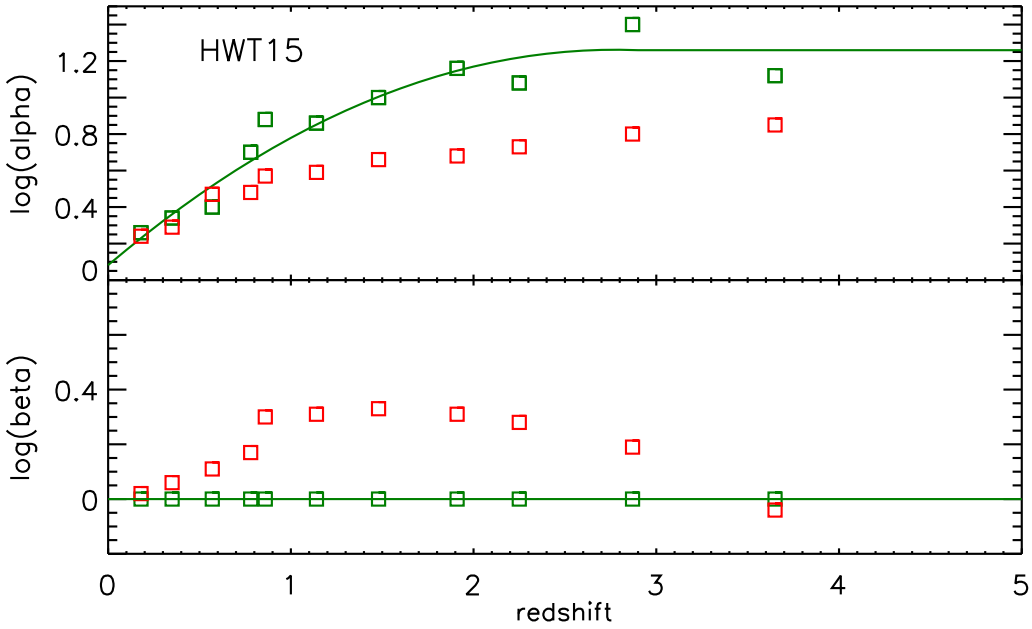

Figure 3.3: evolution of fitting parameters. The top panel shows the value of α, while the bottom panel presents the value of β from fitting to Gruppioni's LF (green symbol) and Bethermin's LF (red symbol).

3.6 Number counts

We are able to produce a mock catalogue directly using the method described in Section 3.3 and present the number counts as well as luminosity function. We also adjust the original individual fluxes by employing the fitted parameters gained from the previous section to observe how the simulated number count is affected. In Section 3.6.1 below, the number counts also were determined by extracting point sources from the simulated map using the approach explained above.

3.6.1 Comparison between the model and the observed counts

Our estimation of the number counts at 250, 350 and 500 micron can be seen in Fig. 3.5. The upper panel shows the number count as a function of flux, where the solid blue line is our original prediction using Equation 3.2 while the dashed lines are those after enhancing the fluxes by applying the corresponding fitted-parameter at each redshift bin. The red and green dash lines are the adjusted number counts using the fitting parameters to Bethermin's LF and Gruppioni's LF where the parameters are taken from Fig. 3.3. Square, triangle and diamond data points come from Béthermin et al. (2012b) and Oliver et al.

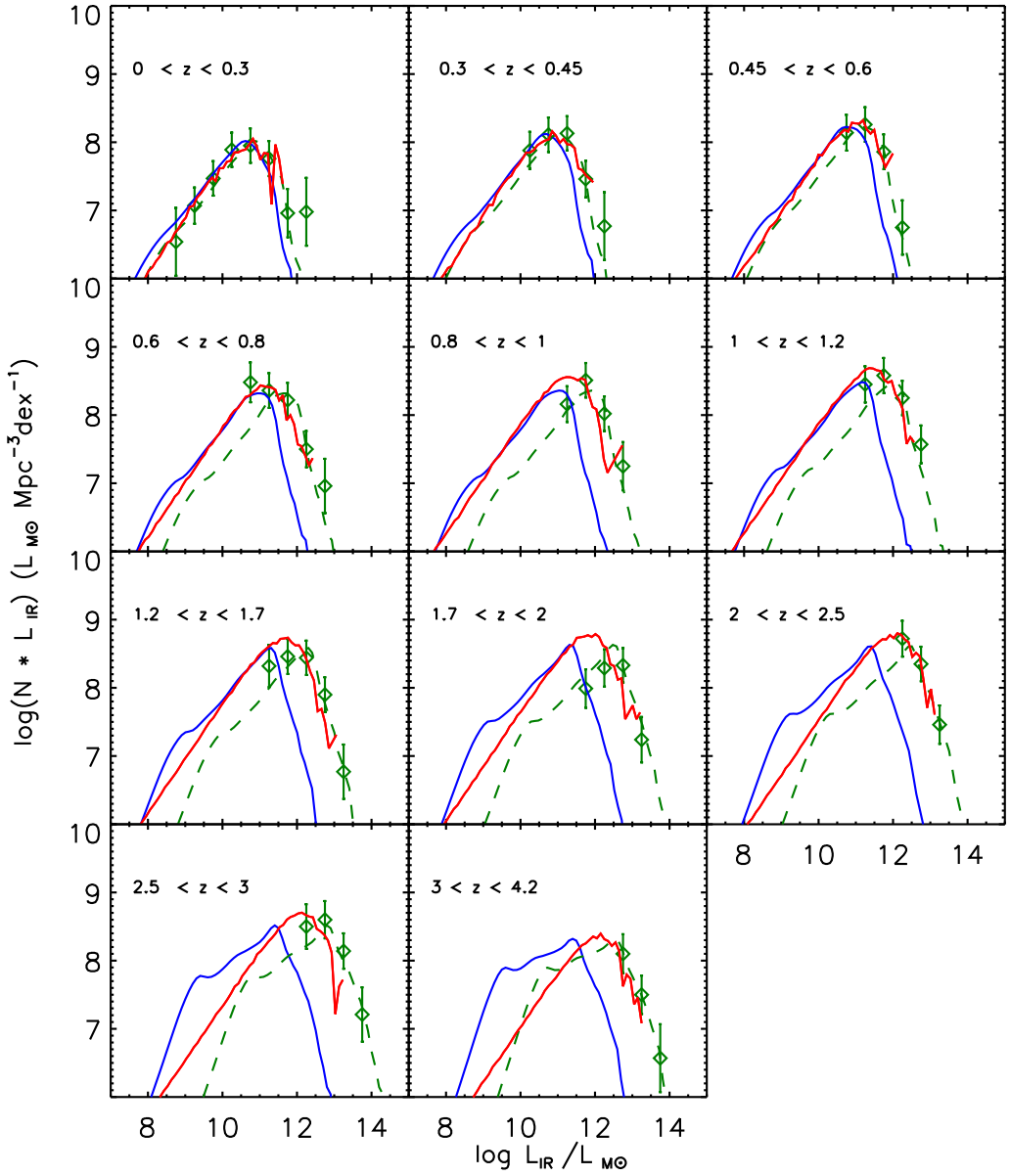

Figure 3.4: The bolometric FIR luminosity function for the original (blue, solid line) and scaled (green, dashed line), as described in the text. Also shown by the solid red line is the Bethermin LFs (Béthermin et al., 2012b). The data points are from Gruppioni et al. (2013).

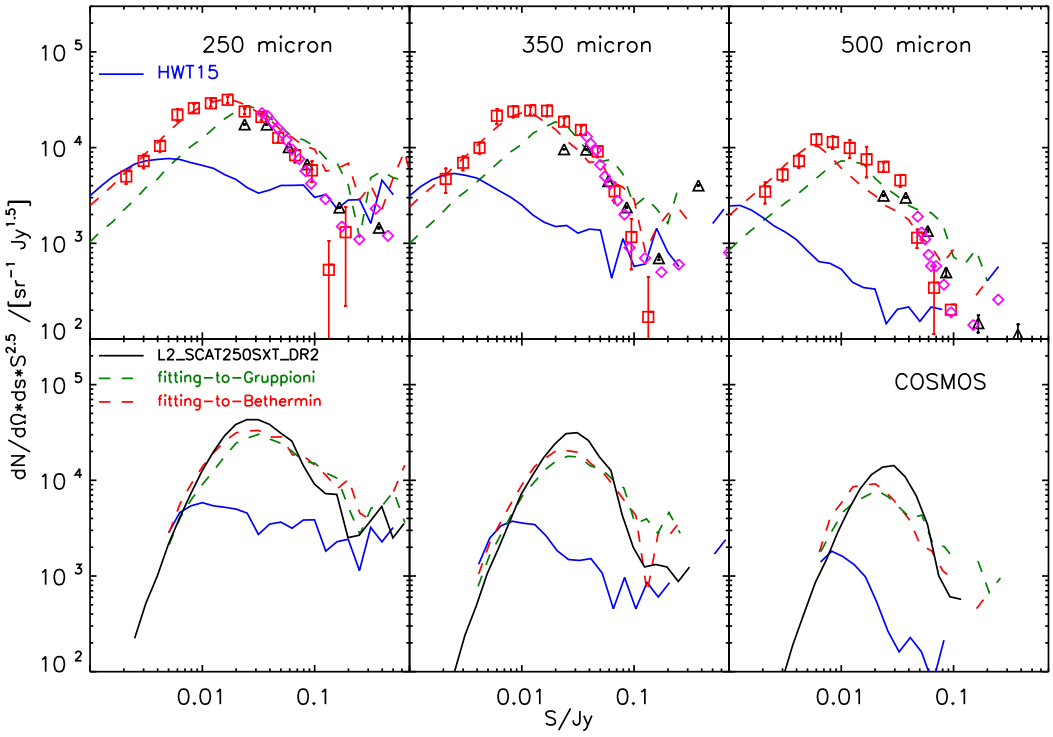

Figure 3.5: Upper panel: 250, 350 and 500 μ number count predicted by the model (solid blue line) and the number count after boosting the fluxes (dashed lines). Symbols with the error bars are from Béthermin et al. (2012b), Oliver et al. (2010) and Clements et al. (2010). Lower panel: Raw number counts as a function of flux by extracting sources from the map using SUSSEXtractor in the COSMOS field. Solid blue line is the result from our mock catalogue and dashed lines are the number count after normalising the fluxes using the parameters from Fig. 3.3. Black line shows the raw observational number count from COSMOS field (Smith et al., 2012).

(2010) and Clements et al. (2010) respectively. The original estimation underpredicts the number density for all but the highest and lowest fluxes. After modifying, the green line now matches the general shape of the observed relation reasonably well and fits better to the data point at the higher end of the number counts, but shifted slightly to higher fluxes. It seems that in spite of overall matching, the most recent observations of total IR LF from z=0.5 to z=4, we predict a very strong increase at the higher end of the 250 number counts and underestimate the lower wing. As a consequence, there is some tension in our prediction that causes even an additional excess of number counts. In contrast, the red dash line matches the data point quite well up to 0.1 Jy, however it is above the data point for the higher fluxes.

As Fig. 3.4 shows, for z<2, Bethermin's LFs basically follow the faint shoulder of our original LFs while at the bright end, they follow our scaled LFs. The deficit of faint objects that is seen in the number counts is undoubtedly a consequence of the fact that our rescaled LF fit lies below the Bethermin LFs at low luminosities. We note that most of the high flux sources come from 0.5<z<1.5 which suggest that the extra number density of bright objects in our scaled number count is due to our higher LF after enhancing the luminosity at these redshift intervals in comparison with Bethermin's.

As described previously we fitted our predicted LFs to those obtained by Gruppioni and used those fitted parameters to scale our number counts. We showed that our boosted number count overpredicts at higher fluxes in comparison with the data. Full explanations of the result are discussed above. We should stress that Gruppioni LFs are not the underpinning luminosities for Bethermin number counts; therefore, fitting our LFs to Gruppioni doesn't necessarily lead the simulated number count to fit Bethermin's data. As a result, when we fit our LF to Bethermin's LF instead of Gruppioni's and don't force our selves to preserve the total flux, a new set of scaled parameters is gained. it can be seen from Fig. 3.5 that by applying these new parameters to the simulated number counts, we will be able to match the number count data points because this time we use the corresponding underpinning LF for Bethermin number counts.

The lower panel of Fig. 3.5 shows the number count from the catalogue produced through map making and source extraction, SMAP and SUSSEXtractor, respectively (Levenson et al., 2010b; Savage & Oliver, 2007). The simulated catalogues were processed by the SMAP simulator to create a mock map with similar properties. The result map was then processed in SUSSEXtractor (SXT), a simple peak finder, in order to provide a point source catalogue. Therefore, these are the raw number counts without any correction

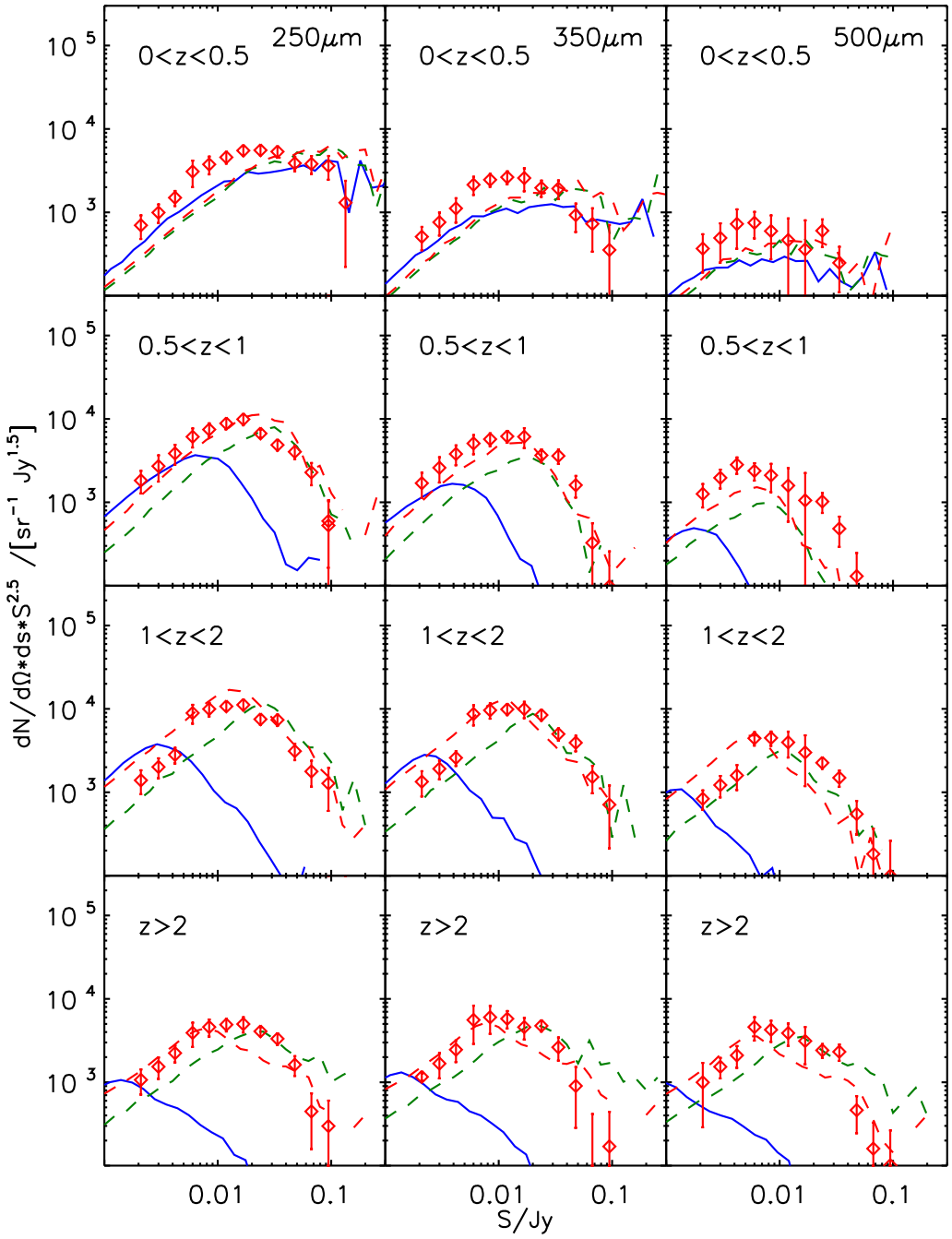

Figure 3.6: Comparison between simulated number counts and observations (Béthermin et al., 2012b) for different redshift intervals. The blue line shows the original model estimation, dash green and red line present the model estimation after boosting the fluxes by applying the corresponding fitted-parameter to Gruppioni and Bethermin LF respectively (Béthermin et al., 2012b; Gruppioni et al., 2013).

for any observational effects. As for the top panel, we present the results from our true prediction and the modified model with the solid green line and the dashed green line, respectively. The observation is a single band catalogue of blind extractions with SUSSEXtractor from the second public data release (DR2) from COSMOS field (Smith et al., 2012). Through the above process, the number counts of the bright fluxes almost remain the same whereas they substantially drop at lower fluxes due to the significant amount of noise and incompleteness. In addition, the confusion noise also scatters the bright fluxes up due to flux boosting which leads to removing the counts from faint fluxes.

In order to better illustrate where there is a shortfall or possibly an excess of predicted fluxes we demonstrate our prediction of the number count before and after fitting to observation's of LFs in different redshift intervals. Fig. 3.6 shows the simulated number count split into four different redshift bins. As we can see from the figure, the original prediction of the number count (blue line) is below the observed data points and this shortage increases strongly with the redshift. The trend of the number counts with redshift is different after scaling the fluxes using the method explained previously. In general, the number density is boosted at the given flux. Both dashed lines are following the shape of the data points quite well. The red line which is obtained from fitting to Bethermin's LF covers the lower tail of the observed data while it grows above the data points at the higher fluxes while the green line that is achieved from fitting to Gruppioni's LF under predicts the number density of faint objects but over predicts it at higher fluxes.The latter one even is lifted above the former for bright objects at the 2 last redshift bins.

3.6.2 $P(D)$

Although in the above sections we made the comparison between the fluxes extracted from our model map and those of the Herschel observed map, we know that this method is affected by confusion and instrument noise. Therefore, it is crucial to have another method to investigate the result with more accuracy and distinguish the objects sourcing the outcome flux in our map. $P(D)$ is simply measuring the pixel distribution in the map which is the number of pixels as a function of the flux density bin.

The shape of $P(D)$ can depend on the shape of the bin and the number count density curve below the detection level where the fluxes are too faint to be resolved, Barcons et al. (1994). Maloney et al. (2005) shows that this approach can provide more precise constrains on the number count density in comparison to those resulting by extracting individual sources. For the reasons mentioned above, this technique is widely applied by

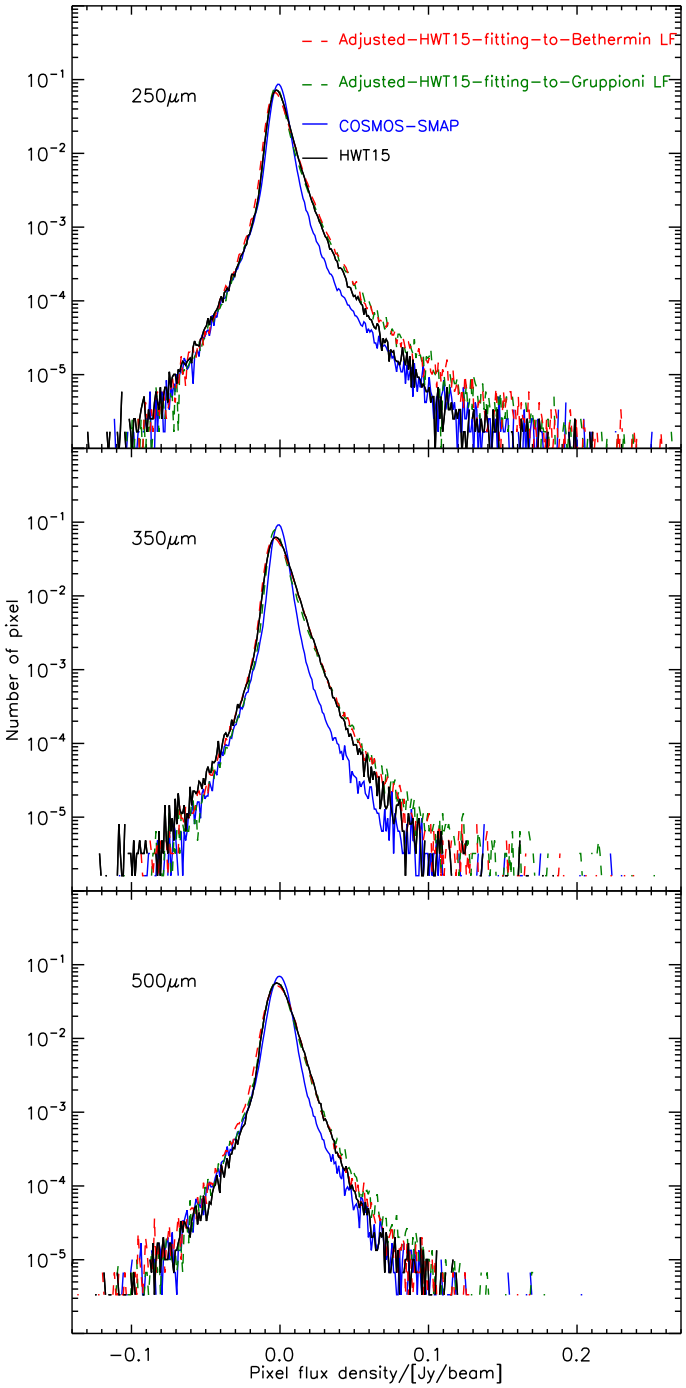

Figure 3.7: Comparison of the pixel flux density between our predicted model and observations for COSMOS at 250μ (Levenson et al., 2010a). The black line represents the observed $P(D)$, the blue line shows the result from the original simulated map and the dash green and red are simulated $P(D)$ after adjustment.

different observational groups (i.g., Maloney et al. (2005), Patanchon et al. (2009) and Glenn et al. (2010)) to extract source number density in different fields. In Fig. 3.7, we compared our model $P(D)$ with the observed one in the COSMOS field. The solid blue line shows our predicted fluctuations while the black line illustrates the observed result from Levenson et al. (2010a). The maps are subtracting by their own mean value so the peak will be approximately zero. It is worth to mention that our purpose is not extracting source density from this technique, however, we are keen to examine our model as to whether it is able to produce correct pixel fluctuations and evaluate our model prediction of number density.

Although both model and observation follow one another on the lower tail of fluctuation, the original model $P(D)$ (before adjustment) shows a deficit between 0.01 and 0.07 Jy, while after scaling (using the method explained above) the fluxes, they show an excess above about 0.3 Jy in comparison to the real map. Generally speaking, the result of the pixel fluctuation plot is consistent with the outcome of Fig. 3.5. The differences between the number counts of the simulated and the real data leave their footprints on the $P(D)$.

3.7 Discussion

In this paper, we have extended the L-Galaxies SAM (Henriques et al., 2015) model to predict the far-infrared number counts and luminosity function. We used the Kennicutt (1998) relation to convert SFR to infrared luminosity. In order to calculate SPIRE fluxes, we scaled our predicted IR luminosities using the Magdis et al. (2012) main-sequence (MS) SEDs that evolve with redshift. We should stress that for the purpose of this work we just use the main-sequence templates from this library. We also investigated the scaled LD and how it affects the number counts. We compared our prediction of IR LF and 250 micron number counts to those presented by Gruppioni et al. (2013), and Béthermin et al. (2012b) and Oliver et al. (2010), respectively. We also produced a map using the fluxes from this paper. Our key results are as follows:

- Our prediction of total flux density shows that although our prediction of total CIB that is within 2σ agreement, we firmly overestimate CIB emission at the flux density cuts above 0.01 Jy. We should note that, uncertainty is large when the observed CIB is measured.

- The SA model shows poor agreement with the observed IR LF from the Gruppioni paper and the disagreement increases gradually with redshift. The model does not

show the evolution as strongly as Gruppioni's and mostly falls bellow the data points for all redshift ranges.

- After scaling the luminosities to fit to Gruppioni's LF, in a way that preserves the total flux, we get a reasonable qualitative fit to the FIR luminosity function. The model prediction now passes through most of the data points especially when log(L)>10.5.

- Bethermin's LFs fit to the Gruppioni data while they agree with our scaled values at high luminosities predict a higher number of sources at low luminosities.

- We adjust the simulated LF, in order to fit Bethermin's LF using two separate parameters for scaling the luminosities and the number densities. the adjusted result agrees very well with Bethermin's LF.

- We extend L-Galaxies, which enables us to estimate fluxes. Our original prediction for 250 micron shows that the number density is underestimated at fluxes above 0.02 Jy, while we overestimate it at fluxes >0.1Jy.

- Whilst scaling the fluxes using parameters obtained from fitting to Gruppioni's LF, considering the total flux is preserved, the alignment between our model and the observation improves. Although our scaled number counts are below the data points at fluxes <0.02, we estimate greater number density for higher fluxes.

- In contrast, when we used the adjusted parameters gained from fitting to Bethermin's LF, the predicted number counts match the data points really well for fluxes <0.03 Jy, however it shows an excess above about 0.03 Jy in comparison to the observed data.

- The comparison shows that the raw number counts extracted from the map are similar to the number counts when corrected for flux- boosting, completeness and other observational biases. The original model strongly shows less number density for all fluxes while the scaled one aligns better with the real data.

- The result of $P(D)$ is consistent with the contrast of number counts. The unscaled simulated map has less number of pixel than the observed map for pixel flux densities above 0.01and peaks above the observed one, whereas the fitted one overestimates the pixel flux density in comparison with the real map but matches the peak of the observed $P(D)$.

Our unscaled model underestimates the high luminosity sources at high redshift. It produces enough flux to give a good fit if we lump that flux into brighter sources. At the same time that then gives a good fit to the raw number counts. However, the modelling of Bethermin prefers a fit with more faint sources.

Overall, L-Galaxies predicts a slight deficit in the total flux at the 2 σ level. It under-estimates the number density of galaxies at most fluxes, but with a slight excess at the bright end. In general, we need to reduce slightly the FIR luminosity of sources at $z < 0.5$ in our model but increase the luminosity of sources at high redshift.

One way of generating more luminosity is keeping the same number of objects but making star formation more bursty. However L-Galaxies is able to produce the same number of stars in the current Universe and matches the stellar mass of the local universe. As a result, we must consider an approach that generates more luminous objects without losing our fit to the current Universe. Some other SAM models, particularly the Durham version (known as GALFORM (Baugh et al., 2005; Lacey et al., 2015)) solve the shortfall of far-infrared LIR by introducing a variable IMF with a higher SFR in starbursts. The GALFORM results however, show that they over-predict the bright number counts and under-predict at 0.01 Jy.

As mentioned above our aim is keeping the number of stars fixed, thus moving to a top-heavy IMF would means that we could have a higher far-IR luminosity for the same amount of star formation. However, there would also be a danger that we make our galaxies too blue and over-produce metals. One way of solving this problem in our model is changing the IMF or perhaps even the SED with redshift. Perhaps another way is compressing a few more stars into the knee of the mass function at z=0. With this approach we might be able to keep the number of stars unchanged but produce more flux at higher redshifts and still fit the stellar mass function at redshift zero.

As a conclusion, SFR and infrared luminosity are one of the key measures of galaxy evolution. Further work can provide a more direct approach to predicting the number density by using a realistic dust model. Considering the implications of variations of IMF in the framework of L-Galaxies is another investigation that can affect metallicity enrichment and SFR. This point will be explored in future works.

Bibliography

Barcons X., Raymont G. B., Warwick R. S., Fabian A. C., Mason K. O., McHardy I., Rowan-Robinson M., 1994, MNRAS, 268, 833 Cited on 72

Barger A. J., Cowie L. L., Sanders D. B., Fulton E., Taniguchi Y., Sato Y., Kawara K., Okuda H., 1998, Nature, 394, 248 Cited on 58

Baugh C. M., Lacey C. G., Frenk C. S., Granato G. L., Silva L., Bressan A., Benson A. J., Cole S., 2005, MNRAS, 356, 1191 Cited on 59, 76

Béthermin M. et al., 2012a, ApJ, 757, L23 Cited on 61

Béthermin M., Dole H., Beelen A., Aussel H., 2010, A&A, 512, A78 Cited on 62, 63

Béthermin M. et al., 2012b, A&A, 542, A58 Cited on 62, 63, 66, 67, 68, 69, 71, 74

Boylan-Kolchin M., Springel V., White S. D. M., Jenkins A., Lemson G., 2009, MNRAS, 398, 1150 Cited on 60

Calzetti D., 2001, NewAR, 45, 601 Cited on 58

Chabrier G., 2003, PASP, 115, 763 Cited on 61

Chapman S. C., Blain A. W., Smail I., Ivison R. J., 2005, ApJ, 622, 772 Cited on 58

Clements D. L. et al., 2010, A&A, 518, L8 Cited on 69, 70

Cole S., Lacey C. G., Baugh C. M., Frenk C. S., 2000, MNRAS, 319, 168 Cited on 59

Coppin K. et al., 2008, MNRAS, 384, 1597 Cited on 58

Crawford T. M., Switzer E. R., Holzapfel W. L., Reichardt C. L., Marrone D. P., Vieira J. D., 2010, ApJ, 718, 513 Cited on 62

Dole H. et al., 2006, A&A, 451, 417 Cited on 58

Fixsen D. J., Dwek E., Mather J. C., Bennett C. L., Shafer R. A., 1998, ApJ, 508, 123 Cited on 58, 62, 63

Fontanot F., Monaco P., Silva L., Grazian A., 2007, MNRAS, 382, 903 Cited on 59

Gispert R., Lagache G., Puget J. L., 2000, A&A, 360, 1 Cited on 58

Glenn J. et al., 2010, MNRAS, 409, 109 Cited on 74

Gonzalez-Perez V., Lacey C. G., Baugh C. M., Frenk C. S., Wilkins S. M., 2013, MNRAS, 429, 1609 Cited on 59

Griffin M. J. et al., 2010, A&A, 518, L3 Cited on 58

Gruppioni C. et al., 2013, MNRAS Cited on 58, 63, 64, 65, 68, 71, 74

Henriques B., White S., Thomas P., Angulo R., Guo Q., Lemson G., Springel V., Overzier R., 2015, MNRAS, 451, 2263 Cited on 59, 60, 74

Hogg D. W., 1999, ArXiv Astrophysics e-prints Cited on 61

Holland W. S., Cunningham C. R., Gear W. K., Jenness T., Laidlaw K., Lightfoot J. F., Robson E. I., 1998, in Procspie., Vol. 3357, Advanced Technology MMW, Radio, and Terahertz Telescopes, Phillips T. G., ed., pp. 305–318 Cited on 58

Kennicutt, Jr. R. C., 1998, ARA&A, 36, 189 Cited on 57, 61, 65, 74

Kriek M., van Dokkum P. G., Labbé I., Franx M., Illingworth G. D., Marchesini D., Quadri R. F., 2009, ApJ, 700, 221 Cited on 61

Lacey C. G., Baugh C. M., Frenk C. S., Benson A. J., 2011, MNRAS, 412, 1828 Cited on 59

Lacey C. G. et al., 2015, ArXiv e-prints Cited on 59, 76

Lagache G., Haffner L. M., Reynolds R. J., Tufte S. L., 2000, A&A, 354, 247 Cited on 62, 63

Lee S.-K., Idzi R., Ferguson H. C., Somerville R. S., Wiklind T., Giavalisco M., 2009, ApJS, 184, 100 Cited on 59

Levenson L. et al., 2010a, MNRAS, 409, 83 Cited on 62, 73, 74

Levenson L. et al., 2010b, MNRAS, 409, 83 Cited on 70

Magdis G. E. et al., 2012, ApJ, 760, 6 Cited on 61, 74

Magnelli B. et al., 2013, A&A, 553, A132 Cited on 58

Maloney P. R. et al., 2005, ApJ, 635, 1044 Cited on 72, 74

Nguyen H. T. et al., 2010, A&A, 518, L5 Cited on 61

Oliver S. et al., 2010 Cited on 62

Oliver S. J. et al., 2010, A&A, 518, L21 Cited on 62, 63, 67, 69, 74

Ott S., 2010, in Astronomical Society of the Pacific Conference Series, Vol. 434, Astronomical Data Analysis Software and Systems XIX, Mizumoto Y., Morita K.-I., Ohishi M., eds., p. 139 Cited on 62

Pascale E. et al., 2009, ApJ, 707, 1740 Cited on 58

Patanchon G. et al., 2009 Cited on 74

Planck Collaboration XVI, 2014, A&A, 571, A16 Cited on 60

Poglitsch A. et al., 2010, A&A, 518, L2 Cited on 58

Puget J.-L., Abergel A., Bernard J.-P., Boulanger F., Burton W. B., Desert F.-X., Hartmann D., 1996, A&A, 308, L5 Cited on 58

Savage R. S., Oliver S., 2007, ApJ, 661, 1339 Cited on 62, 70

Smith A. J. et al., 2012, MNRAS, 419, 377 Cited on 69, 72

Springel V. et al., 2005, Nature, 435, 629 Cited on 60

Valiante E. et al., 2016 Cited on 62

Valtchanov I. et al., 2014, Experimental Astronomy, 37, 207 Cited on 62

Viero M. P. et al., 2013, ApJ, 772, 77 Cited on 62

Chapter 4

Conclusion

Understanding the physics behind galaxy evolution is one the biggest aims in modern astronomy. In the canonical Lambda CDM theory, we know that the dark matter interacts only via gravity and its evolution is now well established from large N-body simulations. However, the evolution of baryons evolving with DM haloes is a complex process and not fully understood in detail yet. Semi-analytic models have been developed to try to understand the physics that gives rise to galaxies that resemble those in the observed universe. They adopt a set of simplified physical recipes for baryonic processes, which are implemented on top of a merger tree of dark matter haloes. The Munich SA model, known as L-Galaxies provides a good fit to lots of the observed quantities such as the mass and luminosity functions of galaxies, the fraction of quenched galaxies, the star formation versus stellar mass relation and many more.

One of the most fundamental ways to probe the physics behind galaxy evolution is to look at the star formation rate as a function of cosmic time. The main motivations of the first paper in this thesis were comparing the predicted star formation histories (SFHs) of galaxies with that of the observed ones and investigating the possible range of SFHs at high redshifts. We have recorded star formation histories (SFHs) in the L-Galaxies in bins of time and contrasted our SFHs to those in the VESPA catalogue obtained from SDSS-DR7 data. We show that reconstructing stellar spectra in post-processing provides a good agreement in comparison to detailed integration of the stellar emission over cosmic time. The SA prediction of SFH presents reasonable qualitative agreement with the observed SFHs of the SDSS Main Galaxy Sample from the VESPA catalogue. We find that the HWT15 model fits better for stars younger than 1 Gyr, while the GWB11 model fits better for older stars. When divided up by mass, both VESPA and the models show a downsizing trend. We also investigate SFH behaviour when divided by colour. The model prediction

shows a pronounced distinction between the population of red and blue galaxies that begins as long as 5 Gyr ago as opposed to only 0.3 Gyr ago for the VESPA observed galaxies. One of the most fundamental ways to probe the physics behind the galaxy evolution is star formation rate as a function of cosmic time. The main motivations of the first paper of this thesis were comparing the predicted star formation histories (SFHs) of galaxies with that of the observed ones and investigating the possible range of SFHs at high redshifts. We have presented the recording of star formation histories (SFHs) in the L-Galaxies in bins of time and contrasted our SFHs to those in the VESPA catalogue obtained from SDSS-DR7 data. We show that reconstructing stellar spectra in post-processing provides a good agreement in comparison to detailed integration of the stellar emission over cosmic time. The SA prediction of SFH presents reasonable qualitative agreement with the observed SFHs of the SDSS Main Galaxy Sample from the VESPA catalogue. We find that the HWT15 model fits better for stars younger than 1 Gyr, while the GWB11 model fits better for older stars. When divided up by mass, VESPA and the models show downsizing trend. We also investigate SFH behaviour when divided by colour. The model prediction shows a pronounced distinction between the population of red and blue galaxies that begins as long as 5 Gyrs ago as opposed to only 0.3 Gyr ago for the VESPA observed galaxies.

In the second paper, I extend the predictions of L-Galaxies into the far infra-red (FIR), using the model prediction of SFR is used as a proxy for FIR luminosity. Therefore, the model is able to predict the number counts from SPIRE bands and the redshift evolution of the FIR luminosity. The predicted 250, 350 and 500 micron count underestimate the number density of individual objects over broad range of fluxes, as shown in Fig. 1.3 in the paper. I estimate the evolution of the FIR LF using the snapshots where its associated redshift is equal to the average of each redshift interval and contrast it with observations from Gruppioni (Gruppioni et al., 2013). The model and the data are a reasonable match at low redshifts but the slight deviation between the two becomes significant at higher redshifts. To try to reconcile the two, I first try a model that conserves flux but that boosts the emission of individual sources by making them more bursty. After scaling, the predicted counts are still low by factor of 2 at faint fluxes and give too many high fluxes in comparison with the observed counts. Next, I adopt a different approach where the model LFs are fitted to the LF from Bethermin (Béthermin et al., 2012) while the number density and luminosity are scaled independently, meaning that not only does the model luminosity slide to the right but also the number density at the given luminosity is lifted on the plot. The scaled number count plot fits that of the observed data point well on

the first half but over-predicts the number densities associated with fluxes above 0.03 Jy. As mentioned previously, our original model prediction (without scaling) underestimates the number density of the galaxies at the given flux with a slight excess for bright sources with fluxes higher than 0.2 Jy. These high flux sources come predominantly from redshifts below 0.5. By looking at Fig. 3.2, it is obvious that in this redshift range, our LF peaks slightly above the observational one.

As a result, we must consider an approach that generates more luminous objects without over-producing stars in the current-day Universe. Our result strongly suggests that one way of solving this problem is using variable IMF or an SED that changes with redshift. As mentioned above, our aim is keeping the number of stars fixed, thus moving to a top-heavy IMF would mean that we could have a higher FIR luminosity for the same amount of star formation. However, there would also be the risk of making our galaxies too blue and overproducing metals. Some other SAM models, particularly Durham versions (known as GALFORM Baugh et al., 2005; Lacey et al., 2015) solve the shortfall of FIR light by introducing a variable IMF with a higher SFR in starbursts. The GALFORM results however, show that they over-predict the bright number counts and under-predict at 0.01 Jy.

In order to better model the FIR luminosity of individual galaxies in future work, the SFH can be used together with a more sophisticated dust model plus (perhaps) variable FIR SEDs meaning that we will be able to calculate the conversion factor that was first introduced by Kennicutt (1998), from the SFH produced by L-Galaxies.

Bibliography

Baugh C. M., Lacey C. G., Frenk C. S., Granato G. L., Silva L., Bressan A., Benson A. J., Cole S., 2005, MNRAS, 356, 1191 Cited on 82

Béthermin M. et al., 2012, A&A, 542, A58 Cited on 81

Gruppioni C. et al., 2013, MNRAS Cited on 81

Kennicutt, Jr. R. C., 1998, ARA&A, 36, 189 Cited on 82

Lacey C. G. et al., 2015, ArXiv e-prints Cited on 82